

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
24 December 2003 (24.12.2003)

PCT

(10) International Publication Number
WO 03/106654 A2

- (51) International Patent Classification⁷: C12N
Diego, CA 92119 (US). ESTEGHLALIAN, Alireza [CA/US]; 9245 Regents Road, Apt. M302, La Jolla, CA 92037 (US).
- (21) International Application Number: PCT/US03/19153
- (22) International Filing Date: 16 June 2003 (16.06.2003) (74) Agent: EINHORN, Gregory, P.; Fish & Richardson P.C., 4350 La Jolla Village Drive, Suite 500, San Diego, CA 92122 (US).
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/389,299 14 June 2002 (14.06.2002) US (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application: US 60/389,299 (CON) Filed on 14 June 2002 (14.06.2002)
- (71) Applicant (*for all designated States except US*): DIVERSA CORPORATION [US/US]; 4955 Directors Place, San Diego, CA 92121 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): STEER, Brian [CA/US]; 1136 Missouri Street, San Diego, CA 92109 (US). CALLEN, Walter [US/US]; 3469 Stetson Avenue, San Diego, CA 92122 (US). HEALEY, Shaun [US/US]; 7405 Charmant Drive, #2427, San Diego, CA 92122 (US). HAZLEWOOD, Geoff [GB/US]; 13041 Caminito Bautizo, San Diego, CA 92130 (US). WU, Di [CN/US]; 11767 Westview Parkway, San Diego, CA 92126 (US). BLUM, David [US/US]; 7057 Cowles Mountain Boulevard, San Diego, CA 92122 (US).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— *without international search report and to be republished upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 03/106654 A2

(54) Title: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND METHODS FOR MAKING AND USING THEM

(57) Abstract: The invention relates to xylanases and to polynucleotides encoding the xylanases. In addition, methods of designing new xylanases and methods of use thereof are also provided. The xylanases have increased activity and stability at increased pH and temperature.

BEST AVAILABLE COPY

XYLANASES, NUCLEIC ACIDS ENCODING THEM AND METHODS FOR MAKING AND USING THEM

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application claims the benefit of priority under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 60/389,299, filed June 14, 2002. The aforementioned application is explicitly incorporated herein by reference in its entirety and for all purposes.

FIELD OF THE INVENTION

10 This invention relates generally to enzymes, polynucleotides encoding the enzymes, the use of such polynucleotides and polypeptides and more specifically to enzymes having xylanase activity, e.g., catalyzing hydrolysis of internal β -1,4-xylosidic linkages or endo- β -1,4-glucanase linkages.

BACKGROUND

15 Xylanases (e.g., endo-1,4-beta-xylanase, EC 3.2.1.8) hydrolyze internal β -1,4-xylosidic linkages in xylan to produce smaller molecular weight xylose and xylo-oligomers. Xylans are polysaccharides formed from 1,4- β -glycoside-linked D-xylopyranoses. Xylanases are of considerable commercial value, being used in the food industry, for baking and fruit and vegetable processing, breakdown of agricultural waste, in the manufacture of animal feed and in pulp and paper production. Xylanases are formed by
20 fungi and bacteria.

 Arabinoxylanase are major non-starch polysaccharides of cereals representing 2.5 – 7.1% w/w depending on variety and growth conditions. The physicochemical properties of this polysaccharide are such that it gives rise to viscous solutions or even gels under oxidative conditions. In addition, arabinoxylans have high water-binding capacity and
25 may have a role in protein foam stability. All of these characteristics present problems for several industries including brewing, baking, animal nutrition and paper manufacturing. In brewing applications, the presence of xylan results in wort filterability and haze formation issues. In baking applications (especially for cookies and crackers), these arabinoxylans create sticky doughs that are difficult to machine and reduce biscuit size. In addition, this
30 carbohydrate is implicated in rapid rehydration of the baked product resulting in loss of crispiness and reduced shelf-life. For monogastric animal feed applications with cereal diets, arabinoxylan is a major contributing factor to viscosity of gut contents and thereby adversely affects the digestibility of the feed and animal growth rate. For ruminant animals, these

polysaccharides represent substantial components of fiber intake and more complete digestion of arabinoxylans would facilitate higher feed conversion efficiencies.

Xylanases are currently used as additives (dough conditioners) in dough processing for the hydrolysis of water soluble arabinoxylan. In baking applications (especially for cookies and crackers), arabinoxylan creates sticky doughs that are difficult to machine and reduce biscuit size. In addition, this carbohydrate is implicated in rapid rehydration of the baked product resulting in loss of crispiness and reduced shelf-life.

The enhancement of xylan digestion in animal feed may improve the availability and digestibility of valuable carbohydrate and protein feed nutrients. For monogastric animal feed applications with cereal diets, arabinoxylan is a major contributing factor to viscosity of gut contents and thereby adversely affects the digestibility of the feed and animal growth rate. For ruminant animals, these polysaccharides represent substantial components of fiber intake and more complete digestion would facilitate higher feed conversion efficiencies. It is desirable for animal feed xylanases to be active in the animal stomach. This requires a feed enzyme to have high activity at 37 °C and at low pH for monogastrics (pH 2-4) and near neutral pH for ruminants (pH 6.5-7). The enzyme should also possess resistance to animal gut xylanases and stability at the higher temperatures involved in feed pelleting. As such, there is a need in the art for xylanase feed additives for monogastric feed with high specific activity, activity at 35-40°C and pH 2-4, half life greater than 30 minutes in SGF and a half-life > 5 minutes at 85°C in formulated state. For ruminant feed, there is a need for xylanase feed additives that have a high specific activity, activity at 35-40°C and pH 6.5-7.0, half life greater than 30 minutes in SRF and stability as a concentrated dry powder.

Xylanases are also used in a number of other applications. For example, xylanases are used in improving the quality and quantity of milk protein production in lactating cows (see, for example, Kung, L., et al, J. Dairy Science, 2000 Jan 83:115-122), increasing the amount of soluble saccharides in the stomach and small intestine of pigs (see, for example, van der Meulen, J. et al, Arch. Tierernahr, 2001 54:101-115), improving late egg production efficiency and egg yields in hens (see, for example, Jaroni, D., et al, Poult. Sci., 1999 June 78:841-847). Additionally, xylanases have been shown to be useful in biobleaching and treatment of chemical pulps (see, for example, U.S. Pat. No. 5,202,249), biobleaching and treatment of wood or paper pulps (see, for example, U.S. Pat. Nos. 5,179,021, 5,116,746, 5,407,827, 5,405,769, 5,395,765, 5,369,024, 5,457,045, 5,434,071,

5,498,534, 5,591,304, 5,645,686, 5,725,732, 5,759,840, 5,834,301, 5,871,730 and 6,057,438) in reducing lignin in wood and modifying wood (see, for example, U.S. Pat. Nos. 5,486,468 and 5,770,012) as flour, dough and bread improvers (see, for example, U.S. Pat. Nos. 5,108,765 and 5,306,633) as feed additives and/or supplements, as set forth above (see, for example, U.S. Pat. Nos. 5,432,074, 5,429,828, 5,612,055, 5,720,971, 5,981,233, 5,948,667, 6,099,844, 6,132,727 and 6,132,716), in manufacturing cellulose solutions (see, for example, U.S. Pat. No. 5,760,211). Detergent compositions having xylanase activity are used for fruit, vegetables and/or mud and clay compounds (see, for example, U.S. Pat. No. 5,786,316).

Xylanases are also useful in a method of use and composition of a carbohydrase and/or a xylanase for the manufacture of an agent for the treatments and/or prophylaxis of coccidiosis. The manufactured agent can be in the form of a cereal-based animal feed. (see, for example, U.S. Pat. No. 5,624,678) Additional uses for xylanases include use in the production of water soluble dietary fiber (see, for example, U.S. Pat. No. 5,622,738), in improving the filterability, separation and production of starch (see, for example, U.S. Pat. Nos. 4,960,705 and 5,023,176), in the beverage industry in improving filterability of wort or beer (see, for example, U.S. Pat. No. 4,746,517), in an enzyme composition for promoting the secretion of milk of livestock and improving the quality of the milk (see, for example, U.S. Pat. No. 4,144,354), in reducing viscosity of plant material (see, for example, U.S. Pat. No. 5,874,274), in increasing viscosity or gel strength of food products such as jam, marmalade, jelly, juice, paste, soup, salsa, etc. (see, for example, U.S. Pat. No. 6,036,981). Xylanases may also be used in hydrolysis of hemicellulose for which it is selective, particularly in the presence of cellulose. Additionally, the cellulase rich retentate is suitable for the hydrolysis of cellulose (see, for example, U.S. Pat. No. 4,725,544).

Various uses of xylanases include the production of ethanol (see, for example, PCT Application Nos. WO0043496 and WO8100857), in transformation of a microbe that produces ethanol (see, for example, PCT Application No. WO99/46362), in production of oenological tannins and enzymatic composition (see, for example, PCT Application No. WO0164830), in stimulating the natural defenses of plants (see, for example, PCT Application No. WO0130161), in production of sugars from hemicellulose substrates (see, for example, PCT Application No. WO9203541), in the cleaning of fruit, vegetables, mud or clay containing soils (see, for example, PCT Application No. WO9613568), in cleaning beer filtration membranes (see, for example, PCT Application No. WO9623579), in a method of killing or inhibiting microbial cells (see, for example, PCT Application No. WO9732480) and in determining the characteristics of process waters from wood pulp

bleaching by using the ratios of two UV absorption measurements and comparing the spectra (see, for example, PCT Application No. WO9840721).

With regard to xylanases used in the paper and pulp industry, xylanases have been isolated from many sources. In particular, see U.S. Patents No. 6,083,733 and
5 6,140,095 and 6,346,407. In particular, it is noted that U.S. Patents No. 6,140,095 addresses alkali-tolerant xylanases. However, it is noted that there remains a need in the art for xylanases to be used in the paper and pulp industry where the enzyme is active in the temperature range of 65°C to 75°C and at a pH of approximately 10. Additionally, an enzyme of the invention useful in the paper and pulp industry would decrease the need for
10 bleaching chemicals, such as chlorine dioxide.

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior
15 invention.

SUMMARY OF THE INVENTION

The invention provides isolated or recombinant nucleic acids comprising a nucleic acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%,
20 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention, e.g., SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ
25 ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69,
30 SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID

NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, 2000, 2050, 2100, 2200, 2250, 2300, 2350, 2400, 2450, 2500, or more residues, encodes at least one polypeptide having a xylanase activity, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection.

Exemplary nucleic acids of the invention also include isolated or recombinant nucleic acids encoding a polypeptide having a sequence as set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132; SEQ ID NO:134; SEQ ID NO:136; SEQ ID NO:138; SEQ ID NO:140; SEQ ID NO:142; SEQ ID NO:144; NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198, SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228, SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:244, SEQ ID NO:246, SEQ ID NO:248, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256, SEQ ID NO:258, SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:268, SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276, SEQ ID NO:278, SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286, SEQ ID NO:288, SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296, SEQ ID NO:298, SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306, SEQ ID NO:308, SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332, SEQ ID

NO:334, SEQ ID NO:336, SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378 or SEQ ID NO:380, and subsequences thereof and variants thereof. In one aspect, the polypeptide has a xylanase activity.

In one aspect, the invention also provides xylanase-encoding nucleic acids with a common novelty in that they are derived from mixed cultures. The invention provides xylanase-encoding nucleic acids isolated from mixed cultures comprising a nucleic acid sequence having at least about 10, 15, 20, 25, 30, 35, 40, 45, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention, e.g., SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID

NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, over a region of at least about 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, or more.

In one aspect, the invention also provides xylanase-encoding nucleic acids with a common novelty in that they are derived from an environmental source, e.g., mixed environmental sources, a bacterial source and/or an archaeal source, see Table 3, below. In one aspect, the invention provides xylanase-encoding nucleic acids isolated from an environmental source, e.g., a mixed environmental source, a bacterial source and/or an archaeal source, comprising a nucleic acid sequence having at least about 10, 15, 20, 25, 30, 35, 40, 45, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention over a region of at least about 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200 or more, residues, wherein the nucleic acid encodes at least one polypeptide having a xylanase activity, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection.

In one aspect, the invention also provides xylanase-encoding nucleic acids with a common novelty in that they are derived from a common glycosidase family, e.g., family 5, 6, 8, 10, 11, 26 or 30, as set forth in Table 5, below.

5 In one aspect, the sequence comparison algorithm is a BLAST version 2.2.2 algorithm where a filtering setting is set to blastall -p blastp -d "nr pataa" -F F, and all other options are set to default.

Another aspect of the invention is an isolated or recombinant nucleic acid including at least 10 consecutive bases of a nucleic acid sequence of the invention, sequences substantially identical thereto, and the sequences complementary thereto.

10 In one aspect, the xylanase activity comprises catalyzing hydrolysis of internal β -1,4-xylosidic linkages. In one aspect, the xylanase activity comprises an endo-1,4-beta-xylanase activity.

In one aspect, the xylanase activity comprises hydrolyzing a xylan to produce a smaller molecular weight xylose and xylo-oligomer. In one aspect, the xylan comprises an
15 arabinoxylan, such as a water soluble arabinoxylan. The water soluble arabinoxylan can comprise a dough or a bread product.

In one aspect, the xylanase activity comprises hydrolyzing polysaccharides comprising 1,4- β -glycoside-linked D-xylopyranoses. In one aspect, the xylanase activity comprises hydrolyzing hemicelluloses. In one aspect, the xylanase activity comprises
20 hydrolyzing hemicelluloses in a wood or paper pulp or a paper product. In one aspect, the invention provides methods for reducing lignin in a wood or wood product comprising contacting the wood or wood product with a polypeptide of the invention.

In one aspect, the xylanase activity comprises catalyzing hydrolysis of xylans in a beverage or a feed or a food product. The feed or food product can comprise a cereal-
25 based animal feed, a wort or a beer, a milk or a milk product, a fruit or a vegetable. In one aspect, the invention provides a food, feed or beverage or a beverage precursor comprising a polypeptide of the invention. The food can be a dough or a bread product. The beverage or a beverage precursor can be a beer or a wort.

In one aspect, the invention provides methods of dough conditioning
30 comprising contacting a dough or a bread product with at least one polypeptide of the invention under conditions sufficient for conditioning the dough. In one aspect, the invention provides methods of beverage production comprising administration of at least one polypeptide of the invention to a beverage or a beverage precursor under conditions sufficient for decreasing the viscosity of the beverage.

In one aspect, the xylanase activity comprises catalyzing hydrolysis of xylans in a cell, e.g., a plant cell or a microbial cell.

In one aspect, the isolated or recombinant nucleic acid encodes a polypeptide having a xylanase activity that is thermostable. The polypeptide can retain a xylanase activity under conditions comprising a temperature range of between about 37°C to about 95°C; between about 55°C to about 85°C, between about 70°C to about 95°C, or, between about 90°C to about 95°C.

In another aspect, the isolated or recombinant nucleic acid encodes a polypeptide having a xylanase activity that is thermotolerant. The polypeptide can retain a xylanase activity after exposure to a temperature in the range from greater than 37°C to about 95°C or anywhere in the range from greater than 55°C to about 85°C. The polypeptide can retain a xylanase activity after exposure to a temperature in the range between about 1°C to about 5°C, between about 5°C to about 15°C, between about 15°C to about 25°C, between about 25°C to about 37°C, between about 37°C to about 95°C, between about 55°C to about 85°C, between about 70°C to about 75°C, or between about 90°C to about 95°C, or more. In one aspect, the polypeptide retains a xylanase activity after exposure to a temperature in the range from greater than 90°C to about 95°C at pH 4.5.

The invention provides isolated or recombinant nucleic acids comprising a sequence that hybridizes under stringent conditions to a nucleic acid comprising a sequence of the invention, e.g., a sequence as set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID

NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, or fragments or subsequences thereof. In one aspect, the nucleic acid encodes a polypeptide having a xylanase activity. The nucleic acid can be at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200 or more residues in length or the full length of the gene or transcript. In one aspect, the stringent conditions include a wash step comprising a wash in 0.2X SSC at a temperature of about 65°C for about 15 minutes.

The invention provides a nucleic acid probe for identifying a nucleic acid encoding a polypeptide having a xylanase activity, wherein the probe comprises at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 or more, consecutive bases of a sequence comprising a sequence of the invention, or fragments or subsequences

thereof, wherein the probe identifies the nucleic acid by binding or hybridization. The probe can comprise an oligonucleotide comprising at least about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, or about 60 to 100 consecutive bases of a sequence comprising a sequence of the invention, or fragments or subsequences thereof.

5 The invention provides a nucleic acid probe for identifying a nucleic acid encoding a polypeptide having a xylanase activity, wherein the probe comprises a nucleic acid comprising a sequence at least about 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 or more residues having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%,
10 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to a nucleic acid of the invention, wherein the sequence identities are determined by analysis with a sequence comparison algorithm or by visual inspection.

15 The probe can comprise an oligonucleotide comprising at least about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, or about 60 to 100 consecutive bases of a nucleic acid sequence of the invention, or a subsequence thereof.

 The invention provides an amplification primer pair for amplifying a nucleic acid encoding a polypeptide having a xylanase activity, wherein the primer pair is capable of
20 amplifying a nucleic acid comprising a sequence of the invention, or fragments or subsequences thereof. One or each member of the amplification primer sequence pair can comprise an oligonucleotide comprising at least about 10 to 50 consecutive bases of the sequence, or about 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more consecutive bases of the sequence.

25 The invention provides amplification primer pairs, wherein the primer pair comprises a first member having a sequence as set forth by about the first (the 5') 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more residues of a nucleic acid of the invention, and a second member having a sequence as set forth by about the first (the 5') 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more residues of
30 the complementary strand of the first member.

 The invention provides xylanase-encoding nucleic acids generated by amplification, e.g., polymerase chain reaction (PCR), using an amplification primer pair of the invention. The invention provides xylanases generated by amplification, e.g., polymerase chain reaction (PCR), using an amplification primer pair of the invention. The invention

provides methods of making a xylanase by amplification, e.g., polymerase chain reaction (PCR), using an amplification primer pair of the invention. In one aspect, the amplification primer pair amplifies a nucleic acid from a library, e.g., a gene library, such as an environmental library.

5 The invention provides methods of amplifying a nucleic acid encoding a polypeptide having a xylanase activity comprising amplification of a template nucleic acid with an amplification primer sequence pair capable of amplifying a nucleic acid sequence of the invention, or fragments or subsequences thereof.

10 The invention provides expression cassettes comprising a nucleic acid of the invention or a subsequence thereof. In one aspect, the expression cassette can comprise the nucleic acid that is operably linked to a promoter. The promoter can be a viral, bacterial, mammalian or plant promoter. In one aspect, the plant promoter can be a potato, rice, corn, wheat, tobacco or barley promoter. The promoter can be a constitutive promoter. The constitutive promoter can comprise CaMV35S. In another aspect, the promoter can be an
15 inducible promoter. In one aspect, the promoter can be a tissue-specific promoter or an environmentally regulated or a developmentally regulated promoter. Thus, the promoter can be, e.g., a seed-specific, a leaf-specific, a root-specific, a stem-specific or an abscission-induced promoter. In one aspect, the expression cassette can further comprise a plant or plant virus expression vector.

20 The invention provides cloning vehicles comprising an expression cassette (e.g., a vector) of the invention or a nucleic acid of the invention. The cloning vehicle can be a viral vector, a plasmid, a phage, a phagemid, a cosmid, a fosmid, a bacteriophage or an artificial chromosome. The viral vector can comprise an adenovirus vector, a retroviral vector or an adeno-associated viral vector. The cloning vehicle can comprise a bacterial
25 artificial chromosome (BAC), a plasmid, a bacteriophage P1-derived vector (PAC), a yeast artificial chromosome (YAC), or a mammalian artificial chromosome (MAC).

 The invention provides transformed cell comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention, or a cloning vehicle of the invention. In one aspect, the transformed cell can be a bacterial cell, a mammalian cell, a
30 fungal cell, a yeast cell, an insect cell or a plant cell. In one aspect, the plant cell can be a cereal, a potato, wheat, rice, corn, tobacco or barley cell.

 The invention provides transgenic non-human animals comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention. In one aspect, the animal is a mouse.

The invention provides transgenic plants comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention. The transgenic plant can be a cereal plant, a corn plant, a potato plant, a tomato plant, a wheat plant, an oilseed plant, a rapeseed plant, a soybean plant, a rice plant, a barley plant or a tobacco plant.

5 The invention provides transgenic seeds comprising a nucleic acid of the invention or an expression cassette (e.g., a vector) of the invention. The transgenic seed can be a cereal plant, a corn seed, a wheat kernel, an oilseed, a rapeseed, a soybean seed, a palm kernel, a sunflower seed, a sesame seed, a peanut or a tobacco plant seed.

10 The invention provides an antisense oligonucleotide comprising a nucleic acid sequence complementary to or capable of hybridizing under stringent conditions to a nucleic acid of the invention. The invention provides methods of inhibiting the translation of a xylanase message in a cell comprising administering to the cell or expressing in the cell an antisense oligonucleotide comprising a nucleic acid sequence complementary to or capable of hybridizing under stringent conditions to a nucleic acid of the invention. In one aspect, the
15 antisense oligonucleotide is between about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, or about 60 to 100 bases in length.

The invention provides methods of inhibiting the translation of a xylanase message in a cell comprising administering to the cell or expressing in the cell an antisense oligonucleotide comprising a nucleic acid sequence complementary to or capable of
20 hybridizing under stringent conditions to a nucleic acid of the invention. The invention provides double-stranded inhibitory RNA (RNAi) molecules comprising a subsequence of a sequence of the invention. In one aspect, the RNAi is about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more duplex nucleotides in length. The invention provides methods of inhibiting the expression of a xylanase in a cell comprising administering to the cell or expressing in the
25 cell a double-stranded inhibitory RNA (iRNA), wherein the RNA comprises a subsequence of a sequence of the invention.

The invention provides an isolated or recombinant polypeptide comprising an amino acid sequence having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%,
30 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary polypeptide or peptide of the invention over a region of at least about 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350 or more residues, or over the full length of the polypeptide, and the sequence identities are determined

by analysis with a sequence comparison algorithm or by a visual inspection. Exemplary polypeptide or peptide sequences of the invention include SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132; SEQ ID NO:134; SEQ ID NO:136; SEQ ID NO:138; SEQ ID NO:140; SEQ ID NO:142; SEQ ID NO:144; NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198, SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228, SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:244, SEQ ID NO:246, SEQ ID NO:248, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256, SEQ ID NO:258, SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:268, SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276, SEQ ID NO:278, SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286, SEQ ID NO:288, SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296, SEQ ID NO:298, SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306, SEQ ID NO:308, SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334,

SEQ ID NO:336, SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378 or SEQ ID NO:380, and subsequences thereof and variants thereof. Exemplary polypeptides also include fragments of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600 or more residues in length, or over the full length of an enzyme. Exemplary polypeptide or peptide sequences of the invention include sequence encoded by a nucleic acid of the invention.

10 Exemplary polypeptide or peptide sequences of the invention include polypeptides or peptides specifically bound by an antibody of the invention. In one aspect, a polypeptide of the invention has at least one xylanase activity.

Another aspect of the invention provides an isolated or recombinant polypeptide or peptide including at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 or more consecutive bases of a polypeptide or peptide sequence of the invention, sequences substantially identical thereto, and the sequences complementary thereto. The peptide can be, e.g., an immunogenic fragment, a motif (e.g., a binding site), a signal sequence, a prepro sequence or an active site.

15

The invention provides isolated or recombinant nucleic acids comprising a sequence encoding a polypeptide having a xylanase activity and a signal sequence, wherein the nucleic acid comprises a sequence of the invention. The signal sequence can be derived from another xylanase or a non-xylanase (a heterologous) enzyme. The invention provides isolated or recombinant nucleic acids comprising a sequence encoding a polypeptide having a xylanase activity, wherein the sequence does not contain a signal sequence and the nucleic acid comprises a sequence of the invention.

20

25

In one aspect, the xylanase activity comprises catalyzing hydrolysis of internal β -1,4-xylosidic linkages. In one aspect, the xylanase activity comprises an endo-1,4-beta-xylanase activity. In one aspect, the xylanase activity comprises hydrolyzing a xylan to produce a smaller molecular weight xylose and xylo-oligomer. In one aspect, the xylan comprises an arabinoxylan, such as a water soluble arabinoxylan. The water soluble arabinoxylan can comprise a dough or a bread product.

30

In one aspect, the xylanase activity comprises hydrolyzing polysaccharides comprising 1,4- β -glycoside-linked D-xylopyranoses. In one aspect, the xylanase activity

comprises hydrolyzing hemicelluloses. In one aspect, the xylanase activity comprises hydrolyzing hemicelluloses in a wood or paper pulp or a paper product.

In one aspect, the xylanase activity comprises catalyzing hydrolysis of xylans in a feed or a food product. The feed or food product can comprise a cereal-based animal
5 feed, a wort or a beer, a milk or a milk product, a fruit or a vegetable.

In one aspect, the xylanase activity comprises catalyzing hydrolysis of xylans in a cell, e.g., a plant cell or a microbial cell.

In one aspect, the xylanase activity is thermostable. The polypeptide can retain a xylanase activity under conditions comprising a temperature range of between about
10 1°C to about 5°C, between about 5°C to about 15°C, between about 15°C to about 25°C, between about 25°C to about 37°C, between about 37°C to about 95°C, between about 55°C to about 85°C, between about 70°C to about 75°C, or between about 90°C to about 95°C, or more. In another aspect, the xylanase activity can be thermotolerant. The polypeptide can retain a xylanase activity after exposure to a temperature in the range from greater than 37°C
15 to about 95°C, or in the range from greater than 55°C to about 85°C. In one aspect, the polypeptide can retain a xylanase activity after exposure to a temperature in the range from greater than 90°C to about 95°C at pH 4.5.

In one aspect, the isolated or recombinant polypeptide can comprise the polypeptide of the invention that lacks a signal sequence. In one aspect, the isolated or
20 recombinant polypeptide can comprise the polypeptide of the invention comprising a heterologous signal sequence, such as a heterologous xylanase or non-xylanase signal sequence.

In one aspect, the invention provides chimeric proteins comprising a first domain comprising a signal sequence of the invention and at least a second domain. The
25 protein can be a fusion protein. The second domain can comprise an enzyme. The enzyme can be a xylanase.

The invention provides chimeric polypeptides comprising at least a first domain comprising signal peptide (SP), a prepro sequence and/or a catalytic domain (CD) of the invention and at least a second domain comprising a heterologous polypeptide or peptide,
30 wherein the heterologous polypeptide or peptide is not naturally associated with the signal peptide (SP), prepro sequence and/or catalytic domain (CD). In one aspect, the heterologous polypeptide or peptide is not a xylanase. The heterologous polypeptide or peptide can be amino terminal to, carboxy terminal to or on both ends of the signal peptide (SP), prepro sequence and/or catalytic domain (CD).

The invention provides isolated or recombinant nucleic acids encoding a chimeric polypeptide, wherein the chimeric polypeptide comprises at least a first domain comprising signal peptide (SP), a prepro domain and/or a catalytic domain (CD) of the invention and at least a second domain comprising a heterologous polypeptide or peptide, wherein the heterologous polypeptide or peptide is not naturally associated with the signal peptide (SP), prepro domain and/ or catalytic domain (CD).

The invention provides isolated or recombinant signal sequences (e.g., signal peptides) consisting of a sequence as set forth in residues 1 to 14, 1 to 15, 1 to 16, 1 to 17, 1 to 18, 1 to 19, 1 to 20, 1 to 21, 1 to 22, 1 to 23, 1 to 24, 1 to 25, 1 to 26, 1 to 27, 1 to 28, 1 to 28, 1 to 30, 1 to 31, 1 to 32, 1 to 33, 1 to 34, 1 to 35, 1 to 36, 1 to 37, 1 to 38, 1 to 40, 1 to 41, 1 to 42, 1 to 43 or 1 to 44, of a polypeptide of the invention, e.g., SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132; SEQ ID NO:134; SEQ ID NO:136; SEQ ID NO:138; SEQ ID NO:140; SEQ ID NO:142; SEQ ID NO:144; NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198, SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228, SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:244,

SEQ ID NO:246, SEQ ID NO:248, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254,
SEQ ID NO:256, SEQ ID NO:258, SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264,
SEQ ID NO:266, SEQ ID NO:268, SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274,
SEQ ID NO:276, SEQ ID NO:278, SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284,
5 SEQ ID NO:286, SEQ ID NO:288, SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294,
SEQ ID NO:296, SEQ ID NO:298, SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304,
SEQ ID NO:306, SEQ ID NO:308, SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314,
SEQ ID NO:316, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:322, SEQ ID NO:324,
SEQ ID NO:326, SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334,
10 SEQ ID NO:336, SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344,
SEQ ID NO:346, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354,
SEQ ID NO:356, SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364,
SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374,
SEQ ID NO:376, SEQ ID NO:378 or SEQ ID NO:380.

15 In one aspect, the xylanase activity comprises a specific activity at about 37°C
in the range from about 1 to about 1200 units per milligram of protein, or, about 100 to about
1000 units per milligram of protein. In another aspect, the xylanase activity comprises a
specific activity from about 100 to about 1000 units per milligram of protein, or, from about
500 to about 750 units per milligram of protein. Alternatively, the xylanase activity
20 comprises a specific activity at 37°C in the range from about 1 to about 750 units per
milligram of protein, or, from about 500 to about 1200 units per milligram of protein. In one
aspect, the xylanase activity comprises a specific activity at 37°C in the range from about 1 to
about 500 units per milligram of protein, or, from about 750 to about 1000 units per
milligram of protein. In another aspect, the xylanase activity comprises a specific activity at
25 37°C in the range from about 1 to about 250 units per milligram of protein. Alternatively, the
xylanase activity comprises a specific activity at 37°C in the range from about 1 to about 100
units per milligram of protein. In another aspect, the thermotolerance comprises retention of
at least half of the specific activity of the xylanase at 37°C after being heated to the elevated
temperature. Alternatively, the thermotolerance can comprise retention of specific activity at
30 37°C in the range from about 1 to about 1200 units per milligram of protein, or, from about
500 to about 1000 units per milligram of protein, after being heated to the elevated
temperature. In another aspect, the thermotolerance can comprise retention of specific
activity at 37°C in the range from about 1 to about 500 units per milligram of protein after
being heated to the elevated temperature.

The invention provides the isolated or recombinant polypeptide of the invention, wherein the polypeptide comprises at least one glycosylation site. In one aspect, glycosylation can be an N-linked glycosylation. In one aspect, the polypeptide can be glycosylated after being expressed in a *P. pastoris* or a *S. pombe*.

5 In one aspect, the polypeptide can retain a xylanase activity under conditions comprising about pH 6.5, pH 6, pH 5.5, pH 5, pH 4.5 or pH 4. In another aspect, the polypeptide can retain a xylanase activity under conditions comprising about pH 7, pH 7.5 pH 8.0, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5 or pH 11. In one aspect, the polypeptide can retain a xylanase activity after exposure to conditions comprising about pH 6.5, pH 6, pH 5.5,
10 pH 5, pH 4.5 or pH 4. In another aspect, the polypeptide can retain a xylanase activity after exposure to conditions comprising about pH 7, pH 7.5 pH 8.0, pH 8.5, pH 9, pH 9.5, pH 10, pH 10.5 or pH 11.

The invention provides protein preparations comprising a polypeptide of the invention, wherein the protein preparation comprises a liquid, a solid or a gel.

15 The invention provides heterodimers comprising a polypeptide of the invention and a second protein or domain. The second member of the heterodimer can be a different phospholipase, a different enzyme or another protein. In one aspect, the second domain can be a polypeptide and the heterodimer can be a fusion protein. In one aspect, the second domain can be an epitope or a tag. In one aspect, the invention provides homodimers
20 comprising a polypeptide of the invention.

The invention provides immobilized polypeptides having a xylanase activity, wherein the polypeptide comprises a polypeptide of the invention, a polypeptide encoded by a nucleic acid of the invention, or a polypeptide comprising a polypeptide of the invention and a second domain. In one aspect, the polypeptide can be immobilized on a cell, a metal, a
25 resin, a polymer, a ceramic, a glass, a microelectrode, a graphitic particle, a bead, a gel, a plate, an array or a capillary tube.

The invention provides arrays comprising an immobilized nucleic acid of the invention. The invention provides arrays comprising an antibody of the invention.

The invention provides isolated or recombinant antibodies that specifically
30 bind to a polypeptide of the invention or to a polypeptide encoded by a nucleic acid of the invention. The antibody can be a monoclonal or a polyclonal antibody. The invention provides hybridomas comprising an antibody of the invention, e.g., an antibody that specifically binds to a polypeptide of the invention or to a polypeptide encoded by a nucleic acid of the invention.

The invention provides method of isolating or identifying a polypeptide having a xylanase activity comprising the steps of: (a) providing an antibody of the invention; (b) providing a sample comprising polypeptides; and (c) contacting the sample of step (b) with the antibody of step (a) under conditions wherein the antibody can specifically bind to the polypeptide, thereby isolating or identifying a polypeptide having a xylanase activity.

The invention provides methods of making an anti-xylanase antibody comprising administering to a non-human animal a nucleic acid of the invention or a polypeptide of the invention or subsequences thereof in an amount sufficient to generate a humoral immune response, thereby making an anti-xylanase antibody. The invention provides methods of making an anti-xylanase immune comprising administering to a non-human animal a nucleic acid of the invention or a polypeptide of the invention or subsequences thereof in an amount sufficient to generate an immune response.

The invention provides methods of producing a recombinant polypeptide comprising the steps of: (a) providing a nucleic acid of the invention operably linked to a promoter; and (b) expressing the nucleic acid of step (a) under conditions that allow expression of the polypeptide, thereby producing a recombinant polypeptide. In one aspect, the method can further comprise transforming a host cell with the nucleic acid of step (a) followed by expressing the nucleic acid of step (a), thereby producing a recombinant polypeptide in a transformed cell.

The invention provides methods for identifying a polypeptide having a xylanase activity comprising the following steps: (a) providing a polypeptide of the invention; or a polypeptide encoded by a nucleic acid of the invention; (b) providing a xylanase substrate; and (c) contacting the polypeptide or a fragment or variant thereof of step (a) with the substrate of step (b) and detecting a decrease in the amount of substrate or an increase in the amount of a reaction product, wherein a decrease in the amount of the substrate or an increase in the amount of the reaction product detects a polypeptide having a xylanase activity.

The invention provides methods for identifying a xylanase substrate comprising the following steps: (a) providing a polypeptide of the invention; or a polypeptide encoded by a nucleic acid of the invention; (b) providing a test substrate; and (c) contacting the polypeptide of step (a) with the test substrate of step (b) and detecting a decrease in the amount of substrate or an increase in the amount of reaction product, wherein a decrease in the amount of the substrate or an increase in the amount of a reaction product identifies the test substrate as a xylanase substrate.

The invention provides methods of determining whether a test compound specifically binds to a polypeptide comprising the following steps: (a) expressing a nucleic acid or a vector comprising the nucleic acid under conditions permissive for translation of the nucleic acid to a polypeptide, wherein the nucleic acid comprises a nucleic acid of the invention, or, providing a polypeptide of the invention; (b) providing a test compound; (c) contacting the polypeptide with the test compound; and (d) determining whether the test compound of step (b) specifically binds to the polypeptide.

The invention provides methods for identifying a modulator of a xylanase activity comprising the following steps: (a) providing a polypeptide of the invention or a polypeptide encoded by a nucleic acid of the invention; (b) providing a test compound; (c) contacting the polypeptide of step (a) with the test compound of step (b) and measuring an activity of the xylanase, wherein a change in the xylanase activity measured in the presence of the test compound compared to the activity in the absence of the test compound provides a determination that the test compound modulates the xylanase activity. In one aspect, the xylanase activity can be measured by providing a xylanase substrate and detecting a decrease in the amount of the substrate or an increase in the amount of a reaction product, or, an increase in the amount of the substrate or a decrease in the amount of a reaction product. A decrease in the amount of the substrate or an increase in the amount of the reaction product with the test compound as compared to the amount of substrate or reaction product without the test compound identifies the test compound as an activator of xylanase activity. An increase in the amount of the substrate or a decrease in the amount of the reaction product with the test compound as compared to the amount of substrate or reaction product without the test compound identifies the test compound as an inhibitor of xylanase activity.

The invention provides computer systems comprising a processor and a data storage device wherein said data storage device has stored thereon a polypeptide sequence or a nucleic acid sequence of the invention (e.g., a polypeptide encoded by a nucleic acid of the invention). In one aspect, the computer system can further comprise a sequence comparison algorithm and a data storage device having at least one reference sequence stored thereon. In another aspect, the sequence comparison algorithm comprises a computer program that indicates polymorphisms. In one aspect, the computer system can further comprise an identifier that identifies one or more features in said sequence. The invention provides computer readable media having stored thereon a polypeptide sequence or a nucleic acid sequence of the invention. The invention provides methods for identifying a feature in a sequence comprising the steps of: (a) reading the sequence using a computer program which

identifies one or more features in a sequence, wherein the sequence comprises a polypeptide sequence or a nucleic acid sequence of the invention; and (b) identifying one or more features in the sequence with the computer program. The invention provides methods for comparing a first sequence to a second sequence comprising the steps of: (a) reading the first sequence and the second sequence through use of a computer program which compares sequences, wherein the first sequence comprises a polypeptide sequence or a nucleic acid sequence of the invention; and (b) determining differences between the first sequence and the second sequence with the computer program. The step of determining differences between the first sequence and the second sequence can further comprise the step of identifying polymorphisms. In one aspect, the method can further comprise an identifier that identifies one or more features in a sequence. In another aspect, the method can comprise reading the first sequence using a computer program and identifying one or more features in the sequence.

The invention provides methods for isolating or recovering a nucleic acid encoding a polypeptide having a xylanase activity from an environmental sample comprising the steps of: (a) providing an amplification primer sequence pair for amplifying a nucleic acid encoding a polypeptide having a xylanase activity, wherein the primer pair is capable of amplifying a nucleic acid of the invention; (b) isolating a nucleic acid from the environmental sample or treating the environmental sample such that nucleic acid in the sample is accessible for hybridization to the amplification primer pair; and, (c) combining the nucleic acid of step (b) with the amplification primer pair of step (a) and amplifying nucleic acid from the environmental sample, thereby isolating or recovering a nucleic acid encoding a polypeptide having a xylanase activity from an environmental sample. One or each member of the amplification primer sequence pair can comprise an oligonucleotide comprising at least about 10 to 50 consecutive bases of a sequence of the invention. In one aspect, the amplification primer sequence pair is an amplification pair of the invention.

The invention provides methods for isolating or recovering a nucleic acid encoding a polypeptide having a xylanase activity from an environmental sample comprising the steps of: (a) providing a polynucleotide probe comprising a nucleic acid of the invention or a subsequence thereof; (b) isolating a nucleic acid from the environmental sample or treating the environmental sample such that nucleic acid in the sample is accessible for hybridization to a polynucleotide probe of step (a); (c) combining the isolated nucleic acid or the treated environmental sample of step (b) with the polynucleotide probe of step (a); and (d) isolating a nucleic acid that specifically hybridizes with the polynucleotide probe of step (a),

thereby isolating or recovering a nucleic acid encoding a polypeptide having a xylanase activity from an environmental sample. The environmental sample can comprise a water sample, a liquid sample, a soil sample, an air sample or a biological sample. In one aspect, the biological sample can be derived from a bacterial cell, a protozoan cell, an insect cell, a yeast cell, a plant cell, a fungal cell or a mammalian cell.

The invention provides methods of generating a variant of a nucleic acid encoding a polypeptide having a xylanase activity comprising the steps of: (a) providing a template nucleic acid comprising a nucleic acid of the invention; and (b) modifying, deleting or adding one or more nucleotides in the template sequence, or a combination thereof, to generate a variant of the template nucleic acid. In one aspect, the method can further comprise expressing the variant nucleic acid to generate a variant xylanase polypeptide. The modifications, additions or deletions can be introduced by a method comprising error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly (e.g., GeneReassembly™, see, e.g., U.S. Patent No. 6,537,776), gene site saturated mutagenesis (GSSM™), synthetic ligation reassembly (SLR) or a combination thereof. In another aspect, the modifications, additions or deletions are introduced by a method comprising recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation and a combination thereof.

In one aspect, the method can be iteratively repeated until a xylanase having an altered or different activity or an altered or different stability from that of a polypeptide encoded by the template nucleic acid is produced. In one aspect, the variant xylanase polypeptide is thermotolerant, and retains some activity after being exposed to an elevated temperature. In another aspect, the variant xylanase polypeptide has increased glycosylation as compared to the xylanase encoded by a template nucleic acid. Alternatively, the variant xylanase polypeptide has a xylanase activity under a high temperature, wherein the xylanase encoded by the template nucleic acid is not active under the high temperature. In one aspect, the method can be iteratively repeated until a xylanase coding sequence having an altered codon usage from that of the template nucleic acid is produced. In another aspect, the

method can be iteratively repeated until a xylanase gene having higher or lower level of message expression or stability from that of the template nucleic acid is produced.

In one aspect, the invention provides isolated or recombinant nucleic acids comprising a sequence as set forth in SEQ ID NO: 189, wherein SEQ ID NO: 189 contains one or more of the following mutations: the nucleotides at positions 22 to 24 are TTC, the nucleotides at positions 31 to 33 are CAC, the nucleotides at positions 34 to 36 are TTG, the nucleotides at positions 49 to 51 are ATA, the nucleotides at positions 31 to 33 are CAT, the nucleotides at positions 67 to 69 are ACG, the nucleotides at positions 178 to 180 are CAC, the nucleotides at positions 190 to 192 are TGT, the nucleotides at positions 190 to 192 are GTA, the nucleotides at positions 190 to 192 are GTT, the nucleotides at positions 193 to 195 are GTG, the nucleotides at positions 202 to 204 are GCT, the nucleotides at positions 235 to 237 are CCA, or the nucleotides at positions 235 to 237 are CCC. In one aspect, the invention provides methods for making a nucleic acid comprising this sequence, wherein the mutations in SEQ ID NO: 189 are obtained by gene site saturated mutagenesis (GSSM™).

In one aspect, the invention provides isolated or recombinant nucleic acids comprising SEQ ID NO: 190, wherein SEQ ID NO: 190 contains one or more of the following mutations: the aspartic acid at amino acid position 8 is phenylalanine, the glutamine at amino acid position 11 is histidine, the asparagine at amino acid position 12 is leucine, the glycine at amino acid position 17 is isoleucine, the threonine at amino acid position 23 is threonine encoded by a codon other than the wild type codon, the glycine at amino acid position 60 is histidine, the proline at amino acid position 64 is cysteine, the proline at amino acid position 64 is valine, the serine at amino acid position 65 is valine, the glycine at amino acid position 68 is isoleucine, the glycine at amino acid position 68 is alanine, or the valine at amino acid position 79 is proline.

The invention provides methods for modifying codons in a nucleic acid encoding a polypeptide having a xylanase activity to increase its expression in a host cell, the method comprising the following steps: (a) providing a nucleic acid of the invention encoding a polypeptide having a xylanase activity; and, (b) identifying a non-preferred or a less preferred codon in the nucleic acid of step (a) and replacing it with a preferred or neutrally used codon encoding the same amino acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to increase its expression in a host cell.

The invention provides methods for modifying codons in a nucleic acid encoding a polypeptide having a xylanase activity; the method comprising the following steps: (a) providing a nucleic acid of the invention; and, (b) identifying a codon in the nucleic acid of step (a) and replacing it with a different codon encoding the same amino acid as the replaced codon, thereby modifying codons in a nucleic acid encoding a xylanase.

The invention provides methods for modifying codons in a nucleic acid encoding a polypeptide having a xylanase activity to increase its expression in a host cell, the method comprising the following steps: (a) providing a nucleic acid of the invention encoding a xylanase polypeptide; and, (b) identifying a non-preferred or a less preferred codon in the nucleic acid of step (a) and replacing it with a preferred or neutrally used codon encoding the same amino acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to increase its expression in a host cell.

The invention provides methods for modifying a codon in a nucleic acid encoding a polypeptide having a xylanase activity to decrease its expression in a host cell, the method comprising the following steps: (a) providing a nucleic acid of the invention; and (b) identifying at least one preferred codon in the nucleic acid of step (a) and replacing it with a non-preferred or less preferred codon encoding the same amino acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in a host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to decrease its expression in a host cell. In one aspect, the host cell can be a bacterial cell, a fungal cell, an insect cell, a yeast cell, a plant cell or a mammalian cell.

The invention provides methods for producing a library of nucleic acids encoding a plurality of modified xylanase active sites or substrate binding sites, wherein the modified active sites or substrate binding sites are derived from a first nucleic acid comprising a sequence encoding a first active site or a first substrate binding site the method comprising the following steps: (a) providing a first nucleic acid encoding a first active site or first substrate binding site, wherein the first nucleic acid sequence comprises a sequence that hybridizes under stringent conditions to a nucleic acid of the invention, and the nucleic acid encodes a xylanase active site or a xylanase substrate binding site; (b) providing a set of mutagenic oligonucleotides that encode naturally-occurring amino acid variants at a plurality of targeted codons in the first nucleic acid; and, (c) using the set of mutagenic

oligonucleotides to generate a set of active site-encoding or substrate binding site-encoding variant nucleic acids encoding a range of amino acid variations at each amino acid codon that was mutagenized, thereby producing a library of nucleic acids encoding a plurality of modified xylanase active sites or substrate binding sites. In one aspect, the method comprises mutagenizing the first nucleic acid of step (a) by a method comprising an optimized directed evolution system, gene site-saturation mutagenesis (GSSM™), synthetic ligation reassembly (SLR), error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly (GeneReassembly™, U.S. Patent No. 6,537,776), gene site saturated mutagenesis (GSSM™), synthetic ligation reassembly (SLR) and a combination thereof. In another aspect, the method comprises mutagenizing the first nucleic acid of step (a) or variants by a method comprising recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation and a combination thereof.

The invention provides methods for making a small molecule comprising the following steps: (a) providing a plurality of biosynthetic enzymes capable of synthesizing or modifying a small molecule, wherein one of the enzymes comprises a xylanase enzyme encoded by a nucleic acid of the invention; (b) providing a substrate for at least one of the enzymes of step (a); and (c) reacting the substrate of step (b) with the enzymes under conditions that facilitate a plurality of biocatalytic reactions to generate a small molecule by a series of biocatalytic reactions. The invention provides methods for modifying a small molecule comprising the following steps: (a) providing a xylanase enzyme, wherein the enzyme comprises a polypeptide of the invention, or, a polypeptide encoded by a nucleic acid of the invention, or a subsequence thereof; (b) providing a small molecule; and (c) reacting the enzyme of step (a) with the small molecule of step (b) under conditions that facilitate an enzymatic reaction catalyzed by the xylanase enzyme, thereby modifying a small molecule by a xylanase enzymatic reaction. In one aspect, the method can comprise a plurality of small molecule substrates for the enzyme of step (a), thereby generating a library of modified small molecules produced by at least one enzymatic reaction catalyzed by the xylanase enzyme. In one aspect, the method can comprise a plurality of additional enzymes under

conditions that facilitate a plurality of biocatalytic reactions by the enzymes to form a library of modified small molecules produced by the plurality of enzymatic reactions. In another aspect, the method can further comprise the step of testing the library to determine if a particular modified small molecule that exhibits a desired activity is present within the library. The step of testing the library can further comprise the steps of systematically eliminating all but one of the biocatalytic reactions used to produce a portion of the plurality of the modified small molecules within the library by testing the portion of the modified small molecule for the presence or absence of the particular modified small molecule with a desired activity, and identifying at least one specific biocatalytic reaction that produces the particular modified small molecule of desired activity.

The invention provides methods for determining a functional fragment of a xylanase enzyme comprising the steps of: (a) providing a xylanase enzyme, wherein the enzyme comprises a polypeptide of the invention, or a polypeptide encoded by a nucleic acid of the invention, or a subsequence thereof; and (b) deleting a plurality of amino acid residues from the sequence of step (a) and testing the remaining subsequence for a xylanase activity, thereby determining a functional fragment of a xylanase enzyme. In one aspect, the xylanase activity is measured by providing a xylanase substrate and detecting a decrease in the amount of the substrate or an increase in the amount of a reaction product.

The invention provides methods for whole cell engineering of new or modified phenotypes by using real-time metabolic flux analysis, the method comprising the following steps: (a) making a modified cell by modifying the genetic composition of a cell, wherein the genetic composition is modified by addition to the cell of a nucleic acid of the invention; (b) culturing the modified cell to generate a plurality of modified cells; (c) measuring at least one metabolic parameter of the cell by monitoring the cell culture of step (b) in real time; and, (d) analyzing the data of step (c) to determine if the measured parameter differs from a comparable measurement in an unmodified cell under similar conditions, thereby identifying an engineered phenotype in the cell using real-time metabolic flux analysis. In one aspect, the genetic composition of the cell can be modified by a method comprising deletion of a sequence or modification of a sequence in the cell, or, knocking out the expression of a gene. In one aspect, the method can further comprise selecting a cell comprising a newly engineered phenotype. In another aspect, the method can comprise culturing the selected cell, thereby generating a new cell strain comprising a newly engineered phenotype.

The invention provides methods of increasing thermotolerance or thermostability of a xylanase polypeptide, the method comprising glycosylating a xylanase polypeptide, wherein the polypeptide comprises at least thirty contiguous amino acids of a polypeptide of the invention; or a polypeptide encoded by a nucleic acid sequence of the invention, thereby increasing the thermotolerance or thermostability of the xylanase polypeptide. In one aspect, the xylanase specific activity can be thermostable or thermotolerant at a temperature in the range from greater than about 37°C to about 95°C.

The invention provides methods for overexpressing a recombinant xylanase polypeptide in a cell comprising expressing a vector comprising a nucleic acid comprising a nucleic acid of the invention or a nucleic acid sequence of the invention, wherein the sequence identities are determined by analysis with a sequence comparison algorithm or by visual inspection, wherein overexpression is effected by use of a high activity promoter, a dicistronic vector or by gene amplification of the vector.

The invention provides methods of making a transgenic plant comprising the following steps: (a) introducing a heterologous nucleic acid sequence into the cell, wherein the heterologous nucleic sequence comprises a nucleic acid sequence of the invention, thereby producing a transformed plant cell; and (b) producing a transgenic plant from the transformed cell. In one aspect, the step (a) can further comprise introducing the heterologous nucleic acid sequence by electroporation or microinjection of plant cell protoplasts. In another aspect, the step (a) can further comprise introducing the heterologous nucleic acid sequence directly to plant tissue by DNA particle bombardment. Alternatively, the step (a) can further comprise introducing the heterologous nucleic acid sequence into the plant cell DNA using an *Agrobacterium tumefaciens* host. In one aspect, the plant cell can be a potato, corn, rice, wheat, tobacco, or barley cell.

The invention provides methods of expressing a heterologous nucleic acid sequence in a plant cell comprising the following steps: (a) transforming the plant cell with a heterologous nucleic acid sequence operably linked to a promoter, wherein the heterologous nucleic sequence comprises a nucleic acid of the invention; (b) growing the plant under conditions wherein the heterologous nucleic acids sequence is expressed in the plant cell.

The invention provides methods of expressing a heterologous nucleic acid sequence in a plant cell comprising the following steps: (a) transforming the plant cell with a heterologous nucleic acid sequence operably linked to a promoter, wherein the heterologous nucleic sequence comprises a sequence of the invention; (b) growing the plant under conditions wherein the heterologous nucleic acids sequence is expressed in the plant cell.

The invention provides methods for hydrolyzing, breaking up or disrupting a xylan-comprising composition comprising the following steps: (a) providing a polypeptide of the invention having a xylanase activity, or a polypeptide encoded by a nucleic acid of the invention; (b) providing a composition comprising a xylan; and (c) contacting the

5 polypeptide of step (a) with the composition of step (b) under conditions wherein the xylanase hydrolyzes, breaks up or disrupts the xylan-comprising composition. In one aspect, the composition comprises a plant cell, a bacterial cell, a yeast cell, an insect cell, or an animal cell. Thus, the composition can comprise any plant or plant part, any xylan-containing food or feed, a waste product and the like. The invention provides methods for
10 liquefying or removing a xylan-comprising composition comprising the following steps: (a) providing a polypeptide of the invention having a xylanase activity, or a polypeptide encoded by a nucleic acid of the invention; (b) providing a composition comprising a xylan; and (c) contacting the polypeptide of step (a) with the composition of step (b) under conditions wherein the xylanase removes, softens or liquefies the xylan-comprising composition.

15 The invention provides detergent compositions comprising a polypeptide of the invention, or a polypeptide encoded by a nucleic acid of the invention, wherein the polypeptide has a xylanase activity. The xylanase can be a nonsurface-active xylanase or a surface-active xylanase. The xylanase can be formulated in a non-aqueous liquid composition, a cast solid, a granular form, a particulate form, a compressed tablet, a gel form,
20 a paste or a slurry form. The invention provides methods for washing an object comprising the following steps: (a) providing a composition comprising a polypeptide of the invention having a xylanase activity, or a polypeptide encoded by a nucleic acid of the invention; (b) providing an object; and (c) contacting the polypeptide of step (a) and the object of step (b) under conditions wherein the composition can wash the object.

25 The invention provides textiles or fabrics, including, e.g., threads, comprising a polypeptide of the invention, or a polypeptide encoded by a nucleic acid of the invention. In one aspect, the textiles or fabrics comprise xylan-containing fibers. The invention provides methods for treating a textile or fabric (e.g., removing a stain from a composition) comprising the following steps: (a) providing a composition comprising a polypeptide of the
30 invention having a xylanase activity, or a polypeptide encoded by a nucleic acid of the invention; (b) providing a textile or fabric comprising a xylan; and (c) contacting the polypeptide of step (a) and the composition of step (b) under conditions wherein the xylanase can treat the textile or fabric (e.g., remove the stain). The invention provides methods for improving the finish of a fabric comprising the following steps: (a) providing a composition

comprising a polypeptide of the invention having a xylanase activity, or a polypeptide encoded by a nucleic acid of the invention; (b) providing a fabric; and (c) contacting the polypeptide of step (a) and the fabric of step (b) under conditions wherein the polypeptide can treat the fabric thereby improving the finish of the fabric. In one aspect, the fabric is a wool or a silk.

The invention provides feeds or foods comprising a polypeptide of the invention, or a polypeptide encoded by a nucleic acid of the invention. The invention provides methods for hydrolyzing xylans in a feed or a food prior to consumption by an animal comprising the following steps: (a) obtaining a feed material comprising a xylanase of the invention, or a xylanase encoded by a nucleic acid of the invention; and (b) adding the polypeptide of step (a) to the feed or food material in an amount sufficient for a sufficient time period to cause hydrolysis of the xylan and formation of a treated food or feed, thereby hydrolyzing the xylans in the food or the feed prior to consumption by the animal. In one aspect, the invention provides methods for hydrolyzing xylans in a feed or a food after consumption by an animal comprising the following steps: (a) obtaining a feed material comprising a xylanase of the invention, or a xylanase encoded by a nucleic acid of the invention; (b) adding the polypeptide of step (a) to the feed or food material; and (c) administering the feed or food material to the animal, wherein after consumption, the xylanase causes hydrolysis of xylans in the feed or food in the digestive tract of the animal. The food or the feed can be, e.g., a cereal, a grain, a corn and the like.

The invention provides food or nutritional supplements for an animal comprising a polypeptide of the invention, e.g., a polypeptide encoded by the nucleic acid of the invention. In one aspect, the polypeptide in the food or nutritional supplement can be glycosylated. The invention provides edible enzyme delivery matrices comprising a polypeptide of the invention, e.g., a polypeptide encoded by the nucleic acid of the invention. In one aspect, the delivery matrix comprises a pellet. In one aspect, the polypeptide can be glycosylated. In one aspect, the xylanase activity is thermotolerant. In another aspect, the xylanase activity is thermostable.

The invention provides a food, a feed or a nutritional supplement comprising a polypeptide of the invention. The invention provides methods for utilizing a xylanase as a nutritional supplement in an animal diet, the method comprising: preparing a nutritional supplement containing a xylanase enzyme comprising at least thirty contiguous amino acids of a polypeptide of the invention; and administering the nutritional supplement to an animal to increase utilization of a xylan contained in a feed or a food ingested by the animal. The

animal can be a human, a ruminant or a monogastric animal. The xylanase enzyme can be prepared by expression of a polynucleotide encoding the xylanase in an organism selected from the group consisting of a bacterium, a yeast, a plant, an insect, a fungus and an animal. The organism can be selected from the group consisting of an *S. pombe*, *S. cerevisiae*, *Pichia pastoris*, *Pseudomonas* sp., *E. coli*, *Streptomyces* sp., *Bacillus* sp. and *Lactobacillus* sp.

The invention provides edible enzyme delivery matrix comprising a thermostable recombinant xylanase enzyme, e.g., a polypeptide of the invention. The invention provides methods for delivering a xylanase supplement to an animal, the method comprising: preparing an edible enzyme delivery matrix in the form of pellets comprising a granulate edible carrier and a thermostable recombinant xylanase enzyme, wherein the pellets readily disperse the xylanase enzyme contained therein into aqueous media, and administering the edible enzyme delivery matrix to the animal. The recombinant xylanase enzyme can comprise a polypeptide of the invention. The granulate edible carrier can comprise a carrier selected from the group consisting of a grain germ, a grain germ that is spent of oil, a hay, an alfalfa, a timothy, a soy hull, a sunflower seed meal and a wheat midd. The edible carrier can comprise grain germ that is spent of oil. The xylanase enzyme can be glycosylated to provide thermostability at pelletizing conditions. The delivery matrix can be formed by pelletizing a mixture comprising a grain germ and a xylanase. The pelletizing conditions can include application of steam. The pelletizing conditions can comprise application of a temperature in excess of about 80°C for about 5 minutes and the enzyme retains a specific activity of at least 350 to about 900 units per milligram of enzyme.

The invention provides methods for improving texture and flavor of a dairy product comprising the following steps: (a) providing a polypeptide of the invention having a xylanase activity, or a xylanase encoded by a nucleic acid of the invention; (b) providing a dairy product; and (c) contacting the polypeptide of step (a) and the dairy product of step (b) under conditions wherein the xylanase can improve the texture or flavor of the dairy product. In one aspect, the dairy product comprises a cheese or a yogurt. The invention provides dairy products comprising a xylanase of the invention, or is encoded by a nucleic acid of the invention.

The invention provides methods for improving the extraction of oil from an oil-rich plant material comprising the following steps: (a) providing a polypeptide of the invention having a xylanase activity, or a xylanase encoded by a nucleic acid of the invention; (b) providing an oil-rich plant material; and (c) contacting the polypeptide of step (a) and the oil-rich plant material. In one aspect, the oil-rich plant material comprises an oil-

rich seed. The oil can be a soybean oil, an olive oil, a rapeseed (canola) oil or a sunflower oil.

The invention provides methods for preparing a fruit or vegetable juice, syrup, puree or extract comprising the following steps: (a) providing a polypeptide of the invention having a xylanase activity, or a xylanase encoded by a nucleic acid of the invention; (b) providing a composition or a liquid comprising a fruit or vegetable material; and (c) contacting the polypeptide of step (a) and the composition, thereby preparing the fruit or vegetable juice, syrup, puree or extract.

The invention provides papers or paper products or paper pulp comprising a xylanase of the invention, or a polypeptide encoded by a nucleic acid of the invention. The invention provides methods for treating a paper or a paper or wood pulp comprising the following steps: (a) providing a polypeptide of the invention having a xylanase activity, or a xylanase encoded by a nucleic acid of the invention; (b) providing a composition comprising a paper or a paper or wood pulp; and (c) contacting the polypeptide of step (a) and the composition of step (b) under conditions wherein the xylanase can treat the paper or paper or wood pulp. In one aspect, the pharmaceutical composition acts as a digestive aid or an anti-microbial (e.g., against *Salmonella*). In one aspect, the treatment is prophylactic. In one aspect, the invention provides oral care products comprising a polypeptide of the invention having a xylanase activity, or a xylanase encoded by a nucleic acid of the invention. The oral care product can comprise a toothpaste, a dental cream, a gel or a tooth powder, an odontic, a mouth wash, a pre- or post brushing rinse formulation, a chewing gum, a lozenge or a candy. The invention provides contact lens cleaning compositions comprising a polypeptide of the invention having a xylanase activity, or a xylanase encoded by a nucleic acid of the invention.

In one aspect, the invention provides methods for eliminating or protecting animals from a microorganism comprising a xylan comprising administering a polypeptide of the invention. The microorganism can be a bacterium comprising a xylan, e.g., *Salmonella*.

The invention provides an isolated nucleic acid having a sequence as set forth in SEQ ID NO:189 and variants thereof having at least 50% sequence identity to SEQ ID NO:189 and encoding polypeptides having xylanase activity. In one aspect, the polypeptide has a xylanase activity, e.g., a thermostable xylanase activity.

The invention provides isolated or recombinant nucleic acids comprising SEQ ID NO:189, wherein SEQ ID NO:189 comprises one or more or all of the following sequence variations: the nucleotides at positions 22 to 24 are TTC, the nucleotides at positions 22 to 24

are TTT, the nucleotides at positions 31 to 33 are CAC, the nucleotides at positions 31 to 33 are CAT, the nucleotides at positions 34 to 36 are TTG, the nucleotides at positions 34 to 36 are TTA, the nucleotides at positions 34 to 36 are CTC, the nucleotides at positions 34 to 36 are CTT, the nucleotides at positions 34 to 36 are CTA, the nucleotides at positions 34 to 36 are CTG, the nucleotides at positions 49 to 51 are ATA, the nucleotides at positions 49 to 51 are ATT, the nucleotides at positions 49 to 51 are ATC, the nucleotides at positions 178 to 180 are CAC, the nucleotides at positions 178 to 180 are CAT, the nucleotides at positions 190 to 192 are TGT, the nucleotides at positions 190 to 192 are TGC, the nucleotides at positions 190 to 192 are GTA, the nucleotides at positions 190 to 192 are GTT, the nucleotides at positions 190 to 192 are GTC, the nucleotides at positions 190 to 192 are GTG, the nucleotides at positions 193 to 195 are GTG, the nucleotides at positions 193 to 195 are GTC, the nucleotides at positions 193 to 195 are GTA, the nucleotides at positions 193 to 195 are GTT, the nucleotides at positions 202 to 204 are ATA, the nucleotides at positions 202 to 204 are ATT, the nucleotides at positions 202 to 204 are ATC, the nucleotides at positions 202 to 204 are GCT, the nucleotides at positions 202 to 204 are GCG, the nucleotides at positions 202 to 204 are GCC, the nucleotides at positions 202 to 204 are GCA, the nucleotides at positions 235 to 237 are CCA, the nucleotides at positions 235 to 237 are CCC, or the nucleotides at positions 235 to 237 are CCG.

The invention provides isolated or recombinant polypeptides comprising an amino acid sequence comprising SEQ ID NO:190, wherein SEQ ID NO:190 comprises one or more or all of the following sequence variations: the aspartic acid at amino acid position 8 is phenylalanine, the glutamine at amino acid position 11 is histidine, the asparagine at amino acid position 12 is leucine, the glycine at amino acid position 17 is isoleucine, the threonine at amino acid position 23 is threonine encoded by a codon other than the wild type codon, the glycine at amino acid position 60 is histidine, the proline at amino acid position 64 is cysteine, the proline at amino acid position 64 is valine, the serine at amino acid position 65 is valine, the glycine at amino acid position 68 is isoleucine, the glycine at amino acid position 68 is alanine, or the serine at amino acid position 79 is proline. In one aspect, the polypeptide has a xylanase activity, e.g., a thermostable xylanase activity.

The invention provides isolated or recombinant nucleic acids comprising SEQ ID NO: 189, wherein SEQ ID NO:189 comprises one or more or all sequence variations set forth in Table 1 or Table 2. The invention provides isolated or recombinant polypeptides encoded by nucleic acids comprising SEQ ID NO: 189, wherein SEQ ID NO:189 comprises

one or more or all sequence variations set forth in Table 1 or Table 2. In one aspect, the polypeptide has a xylanase activity, e.g., a thermostable xylanase activity.

The invention provides isolated or recombinant nucleic acids comprising SEQ ID NO:379, wherein SEQ ID NO:379 comprises one or more or all of the following sequence variations: the nucleotides at positions 22 to 24 are TTC, the nucleotides at positions 31 to 33 are CAC, the nucleotides at positions 49 to 51 are ATA, the nucleotides at positions 178 to 180 are CAC, the nucleotides at positions 193 to 195 are GTG, the nucleotides at positions 202 to 204 are GCT.

The invention provides isolated or recombinant polypeptides comprising SEQ ID NO:380, wherein SEQ ID NO:380 comprises one or more or all of the following sequence variations: D8F, Q11H, G17I, G60H, S65V and/or G68A. In one aspect, the polypeptide has a xylanase activity, e.g., a thermostable xylanase activity.

The isolated or recombinant nucleic acids of the invention are also referred to as "Group A nucleic acid sequences". The invention provides an isolated nucleic acid including at least 10 consecutive bases of a sequence as set forth in Group A nucleic acid sequences, sequences substantially identical thereto and the sequences complementary thereto.

The isolated or recombinant polypeptides of the invention, which include functional fragments of the exemplary sequences of the invention, are also referred to as "Group B amino acid sequences". Another aspect of the invention is an isolated or recombinant nucleic acid encoding a polypeptide having at least 10 consecutive amino acids of a sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto. In yet another aspect, the invention provides a purified polypeptide having a sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto. Another aspect of the invention is an isolated or purified antibody that specifically binds to a polypeptide having a sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto.

Another aspect of the invention is an isolated or purified antibody or binding fragment thereof, which specifically binds to a polypeptide having at least 10 consecutive amino acids of one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto.

Another aspect of the invention is a method of making a polypeptide having a sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto. The method includes introducing a nucleic acid encoding the polypeptide into a host

cell, wherein the nucleic acid is operably linked to a promoter and culturing the host cell under conditions that allow expression of the nucleic acid. Another aspect of the invention is a method of making a polypeptide having at least 10 amino acids of a sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto. The method includes introducing a nucleic acid encoding the polypeptide into a host cell, wherein the nucleic acid is operably linked to a promoter and culturing the host cell under conditions that allow expression of the nucleic acid, thereby producing the polypeptide.

Another aspect of the invention is a method of generating a variant including obtaining a nucleic acid having a sequence as set forth in Group A nucleic acid sequences, sequences substantially identical thereto, sequences complementary to the sequences of Group A nucleic acid sequences, fragments comprising at least 30 consecutive nucleotides of the foregoing sequences and changing one or more nucleotides in the sequence to another nucleotide, deleting one or more nucleotides in the sequence, or adding one or more nucleotides to the sequence.

Another aspect of the invention is a computer readable medium having stored thereon a sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto.

Another aspect of the invention is a computer system including a processor and a data storage device wherein the data storage device has stored thereon a sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide having a sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto.

Another aspect of the invention is a method for comparing a first sequence to a reference sequence wherein the first sequence is a nucleic acid having a sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide code of Group B amino acid sequences and sequences substantially identical thereto. The method includes reading the first sequence and the reference sequence through use of a computer program that compares sequences; and determining differences between the first sequence and the reference sequence with the computer program.

Another aspect of the invention is a method for identifying a feature in a sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide having a sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, including reading the sequence through the use of a

computer program which identifies features in sequences; and identifying features in the sequence with the computer program.

Yet another aspect of the invention is a method of catalyzing the breakdown of xylan or a derivative thereof, comprising the step of contacting a sample containing xylan or the derivative thereof with a polypeptide of Group B amino acid sequences and sequences substantially identical thereto under conditions which facilitate the breakdown of the xylan.

Another aspect of the invention is an assay for identifying fragments or variants of Group B amino acid sequences and sequences substantially identical thereto, which retain the enzymatic function of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto. The assay includes contacting the polypeptide of Group B amino acid sequences, sequences substantially identical thereto, or polypeptide fragment or variant with a substrate molecule under conditions which allow the polypeptide fragment or variant to function and detecting either a decrease in the level of substrate or an increase in the level of the specific reaction product of the reaction between the polypeptide and substrate thereby identifying a fragment or variant of such sequences.

Another aspect of the invention is a nucleic acid probe of an oligonucleotide from about 10 to 50 nucleotides in length and having a segment of at least 10 contiguous nucleotides that is at least 50% complementary to a nucleic acid target region of a nucleic acid sequence selected from the group consisting of Group A nucleic acid sequences; and which hybridizes to the nucleic acid target region under moderate to highly stringent conditions to form a detectable target:probe duplex.

Another aspect of the invention is a polynucleotide probe for isolation or identification of xylanase genes having a sequence which is the same as, or fully complementary to at least a fragment of one of Group A nucleic acid sequences.

In still another aspect, the invention provides a protein preparation comprising a polypeptide having an amino acid sequence selected from Group B amino acid sequences and sequences substantially identical thereto wherein the protein preparation is a liquid.

Still another aspect of the invention provides a protein preparation comprising a polypeptide having an amino acid sequence selected from Group B amino acid sequences and sequences substantially identical thereto wherein the polypeptide is a solid.

Yet another aspect of the invention provides a method for modifying small molecules, comprising the step of mixing at least one polypeptide encoded by a polynucleotide selected from Group A nucleic acid sequences, sequences substantially identical thereto and the sequences complementary thereto with at least one small molecule,

to produce at least one modified small molecule via at least one biocatalytic reaction, where the at least one polypeptide has xylanase activity.

Another aspect of the invention is a cloning vector of a sequence that encodes a polypeptide having xylanase activity, said sequence being selected from Group A nucleic acid sequences, sequences substantially identical thereto and the sequences complementary thereto.

Another aspect of the invention is a host cell comprising a sequence that encodes a polypeptide having xylanase activity, said sequence being selected from Group A nucleic acid sequences, sequences substantially identical thereto and the sequences complementary thereto.

In yet another aspect, the invention provides an expression vector capable of replicating in a host cell comprising a polynucleotide having a sequence selected Group A nucleic acid sequences, sequences substantially identical thereto, sequences complementary thereto and isolated nucleic acids that hybridize to nucleic acids having any of the foregoing sequences under conditions of low, moderate and high stringency.

In another aspect, the invention provides a method of dough conditioning comprising contacting dough with at least one polypeptide of Group B amino acid sequences and sequences substantially identical thereto under conditions sufficient for conditioning the dough.

Another aspect of the invention is a method of beverage production comprising administration of at least one polypeptide of Group B amino acid sequences and sequences substantially identical thereto under conditions sufficient for decreasing the viscosity of wort or beer.

The xylanases of the invention are used to break down the high molecular weight arabinoxylans in animal feed. Adding the xylanases of the invention stimulates growth rates by improving digestibility, which also improves the quality of the animal litter. Xylanase functions through the gastro-intestinal tract to reduce intestinal viscosity and increase diffusion of pancreatic enzymes. Additionally, the xylanases of the invention may be used in the treatment of endosperm cell walls of feed grains and vegetable proteins. In one aspect of the invention, the novel xylanases of the invention are administered to an animal in order to increase the utilization of the xylan in the food. This activity of the xylanases of the invention may be used to break down insoluble cell wall material, liberating nutrients in the cell walls, which then become available to the animal. It also changes hemicellulose to

nutritive sugars so that nutrients formerly trapped within the cell walls are released. Xylanase also produces compounds that may be a nutritive source for the ruminal microflora.

Another aspect of the invention provides a method for utilizing xylanase as a nutritional supplement in the diets of animals, comprising preparation of a nutritional supplement containing a recombinant xylanase enzyme comprising at least thirty contiguous amino acids of Group B amino acid sequences and sequences substantially identical thereto and administering the nutritional supplement to an animal to increase the utilization of xylan contained in food ingested by the animal.

In another aspect of the invention, a method for delivering a xylanase supplement to an animal is provided, where the method comprises preparing an edible enzyme delivery matrix in the form of pellets comprising a granulate edible carrier and a thermostable recombinant xylanase enzyme, wherein the particles readily disperse the xylanase enzyme contained therein into aqueous media, and administering the edible enzyme delivery matrix to the animal. The granulate edible carrier may comprise a carrier selected from the group consisting of grain germ that is spent of oil, hay, alfalfa, timothy, soy hull, sunflower seed meal and wheat midd. The xylanase enzyme may have an amino acid sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto.

In another aspect, the invention provides an isolated nucleic acid comprising a sequence that encodes a polypeptide having xylanase activity, wherein the sequence is selected from Group A nucleic acid sequences, sequences substantially identical thereto and the sequences complementary thereto, wherein the sequence contains a signal sequence. The invention also provides an isolated nucleic acid comprising a sequence that encodes a polypeptide having xylanase activity, wherein the sequence is selected from Group A nucleic acid sequences, sequences substantially identical thereto and the sequences complementary thereto, wherein the sequence contains a signal sequence from another xylanase.

Additionally, the invention provides an isolated nucleic acid comprising a sequence that encodes a polypeptide having xylanase activity, wherein the sequence is selected from Group A nucleic acid sequences, sequences substantially identical thereto and the sequences complementary thereto wherein the sequence does not contain a signal sequence.

Still another aspect of the invention provides an isolated nucleic acid that is a mutation of SEQ ID NO: 189. Yet another aspect provides an amino acid sequence that is a mutation of SEQ ID NO: 190.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

All publications, patents, patent applications, GenBank sequences and ATCC
5 deposits, cited herein are hereby expressly incorporated by reference for all purposes.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings are illustrative of aspects of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

The patent or application file contains at least one drawing executed in color.

10 Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

Figure 1 is a block diagram of a computer system.

Figure 2 is a flow diagram illustrating one aspect of a process for comparing a new nucleotide or protein sequence with a database of sequences in order to determine the
15 homology levels between the new sequence and the sequences in the database.

Figure 3 is a flow diagram illustrating one aspect of a process in a computer for determining whether two sequences are homologous.

Figure 4 is a flow diagram illustrating one aspect of an identifier process 300 for detecting the presence of a feature in a sequence.

20 Figure 5 is a graph comparing activity of the wild type sequence (SEQ ID NOS: 189 and 190) to the 8x mutant (SEQ ID NOS:375, 376), a combination of mutants D, F, H, I, S, V, X and AA in Table 1.

Figure 6A illustrates the nine single site amino acid mutants of SEQ ID NO:378 (encoded by SEQ ID NO:377) as generated by Gene Site Saturation Mutagenesis (GSSM™) of SEQ ID NO:190 (encoded by SEQ ID NO:189), as described in detail in
25 Example 5, below.

Figure 6B illustrates the unfolding of SEQ ID NO:190 and SEQ ID NO:378 in melting temperature transition midpoint (T_m) experiments as determined by DSC for each enzyme, as described in detail in Example 5, below.

30 Figure 7A illustrates the pH and temperature activity profiles for the enzymes SEQ ID NO:190 and SEQ ID NO:378, as described in detail in Example 5, below.

Figure 7B illustrates the rate/temperature activity optima for the enzymes SEQ ID NO:190 and SEQ ID NO:378, as described in detail in Example 5, below.

Figure 7C illustrates the thermal tolerance/ residual activity for the enzymes SEQ ID NO:190 and SEQ ID NO:378, as described in detail in Example 5, below.

Figure 8A illustrates the GeneReassembly™ library of all possible combinations of the 9 GSSM™ point mutations that was constructed and screened for variants with improved thermal tolerance and activity, as described in detail in Example 5, below.

Figure 8B illustrates the relative activity of the “6X-2” variant and “9X” variant (SEQ ID NO:378) compared to SEQ ID NO:190 (“wild-type”) at a temperature optimum and pH 6.0, as described in detail in Example 5, below.

Figure 9A illustrates the fingerprints obtained after hydrolysis of oligoxylans (Xyl)3, (Xyl)4, (Xyl)5 and (Xyl)6 by the SEQ ID NO:190 (“wild-type”) and the “9X” variant (SEQ ID NO:378) enzymes, as described in detail in Example 5, below.

Figure 9B illustrates the fingerprints obtained after hydrolysis of Beechwood xylan by the SEQ ID NO:190 (“wild-type”) and the “9X” variant (SEQ ID NO:378) enzymes, as described in detail in Example 5, below.

Figure 10A is a schematic diagram illustrating the level of thermal stability (represented by Tm) improvement over SEQ ID NO:190 (“wild-type”) obtained by GSSM™ evolution, as described in detail in Example 5, below.

Figure 10B illustrates a “fitness diagram” of enzyme improvement in the form of SEQ ID NO:378 and SEQ ID NO:380, as obtained by combining GSSM™ and GeneReassembly™ technologies, as described in detail in Example 5, below.

Figure 11 is a schematic flow diagram of an exemplary routine screening protocol to determine whether a xylanase of the invention is useful in pretreating paper pulp, as described in detail in Example 6, below.

Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to xylanases and polynucleotides encoding them and methods of making and using them. Xylanase activity of the polypeptides of the invention encompasses enzymes having hydrolase activity, for example, enzymes capable of hydrolyzing glycosidic linkages present in xylan, e.g., catalyzing hydrolysis of internal β -1,4-xylosidic linkages. The xylanases of the invention can be used to make and/or process foods, feeds, nutritional supplements, textiles, detergents and the like. The xylanases of the

invention can be used in pharmaceutical compositions and dietary aids. Xylanases of the invention are particularly useful in baking, animal feed, beverage and paper processes.

Definitions

The term "antibody" includes a peptide or polypeptide derived from, modeled
5 after or substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof, capable of specifically binding an antigen or epitope, see, e.g. Fundamental Immunology, Third Edition, W.E. Paul, ed., Raven Press, N.Y. (1993); Wilson (1994) J. Immunol. Methods 175:267-273; Yarmush (1992) J. Biochem. Biophys. Methods 25:85-97. The term antibody includes antigen-binding portions, i.e., "antigen binding sites,"
10 (e.g., fragments, subsequences, complementarity determining regions (CDRs)) that retain capacity to bind antigen, including (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a F(ab')₂ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the VH and CH1 domains; (iv) a Fv fragment consisting of the VL and VH
15 domains of a single arm of an antibody, (v) a dAb fragment (Ward et al., (1989) Nature 341:544-546), which consists of a VH domain; and (vi) an isolated complementarity determining region (CDR). Single chain antibodies are also included by reference in the term "antibody."

The terms "array" or "microarray" or "biochip" or "chip" as used herein is a
20 plurality of target elements, each target element comprising a defined amount of one or more polypeptides (including antibodies) or nucleic acids immobilized onto a defined area of a substrate surface, as discussed in further detail, below.

As used herein, the terms "computer," "computer program" and "processor" are used in their broadest general contexts and incorporate all such devices, as described in
25 detail, below. A "coding sequence of" or a "sequence encodes" a particular polypeptide or protein, is a nucleic acid sequence which is transcribed and translated into a polypeptide or protein when placed under the control of appropriate regulatory sequences.

The phrases "nucleic acid" or "nucleic acid sequence" as used herein refer to an oligonucleotide, nucleotide, polynucleotide, or to a fragment of any of these, to DNA or
30 RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent a sense or antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material, natural or synthetic in origin. The phrases "nucleic acid" or "nucleic acid sequence" includes oligonucleotide, nucleotide, polynucleotide, or to a fragment of any of these, to DNA or RNA (e.g., mRNA, rRNA, tRNA, iRNA) of genomic or synthetic origin

which may be single-stranded or double-stranded and may represent a sense or antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material, natural or synthetic in origin, including, e.g., iRNA, ribonucleoproteins (e.g., e.g., double stranded iRNAs, e.g., iRNPs). The term encompasses nucleic acids, i.e., oligonucleotides, containing known analogues of natural nucleotides. The term also encompasses nucleic-acid-like structures with synthetic backbones, see e.g., Mata (1997) Toxicol. Appl. Pharmacol. 144:189-197; Strauss-Soukup (1997) Biochemistry 36:8692-8698; Samstag (1996) Antisense Nucleic Acid Drug Dev 6:153-156. "Oligonucleotide" includes either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands that may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide can ligate to a fragment that has not been dephosphorylated.

A "coding sequence of" or a "nucleotide sequence encoding" a particular polypeptide or protein, is a nucleic acid sequence which is transcribed and translated into a polypeptide or protein when placed under the control of appropriate regulatory sequences.

The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as, where applicable, intervening sequences (introns) between individual coding segments (exons). "Operably linked" as used herein refers to a functional relationship between two or more nucleic acid (e.g., DNA) segments. Typically, it refers to the functional relationship of transcriptional regulatory sequence to a transcribed sequence. For example, a promoter is operably linked to a coding sequence, such as a nucleic acid of the invention, if it stimulates or modulates the transcription of the coding sequence in an appropriate host cell or other expression system. Generally, promoter transcriptional regulatory sequences that are operably linked to a transcribed sequence are physically contiguous to the transcribed sequence, i.e., they are cis-acting. However, some transcriptional regulatory sequences, such as enhancers, need not be physically contiguous or located in close proximity to the coding sequences whose transcription they enhance.

The term "expression cassette" as used herein refers to a nucleotide sequence which is capable of affecting expression of a structural gene (i.e., a protein coding sequence, such as a xylanase of the invention) in a host compatible with such sequences. Expression cassettes include at least a promoter operably linked with the polypeptide coding sequence; and, optionally, with other sequences, e.g., transcription termination signals. Additional

factors necessary or helpful in effecting expression may also be used, e.g., enhancers. Thus, expression cassettes also include plasmids, expression vectors, recombinant viruses, any form of recombinant "naked DNA" vector, and the like. A "vector" comprises a nucleic acid that can infect, transfect, transiently or permanently transduce a cell. It will be recognized that a
5 vector can be a naked nucleic acid, or a nucleic acid complexed with protein or lipid. The vector optionally comprises viral or bacterial nucleic acids and/or proteins, and/or membranes (e.g., a cell membrane, a viral lipid envelope, etc.). Vectors include, but are not limited to replicons (e.g., RNA replicons, bacteriophages) to which fragments of DNA may be attached and become replicated. Vectors thus include, but are not limited to RNA, autonomous self-
10 replicating circular or linear DNA or RNA (e.g., plasmids, viruses, and the like, see, e.g., U.S. Patent No. 5,217,879), and include both the expression and non-expression plasmids. Where a recombinant microorganism or cell culture is described as hosting an "expression vector" this includes both extra-chromosomal circular and linear DNA and DNA that has been incorporated into the host chromosome(s). Where a vector is being maintained by a
15 host cell, the vector may either be stably replicated by the cells during mitosis as an autonomous structure, or is incorporated within the host's genome.

As used herein, the term "promoter" includes all sequences capable of driving transcription of a coding sequence in a cell, e.g., a plant cell. Thus, promoters used in the constructs of the invention include *cis*-acting transcriptional control elements and regulatory
20 sequences that are involved in regulating or modulating the timing and/or rate of transcription of a gene. For example, a promoter can be a *cis*-acting transcriptional control element, including an enhancer, a promoter, a transcription terminator, an origin of replication, a chromosomal integration sequence, 5' and 3' untranslated regions, or an intronic sequence, which are involved in transcriptional regulation. These *cis*-acting sequences typically interact
25 with proteins or other biomolecules to carry out (turn on/off, regulate, modulate, etc.) transcription. "Constitutive" promoters are those that drive expression continuously under most environmental conditions and states of development or cell differentiation. "Inducible" or "regulatable" promoters direct expression of the nucleic acid of the invention under the influence of environmental conditions or developmental conditions. Examples of
30 environmental conditions that may affect transcription by inducible promoters include anaerobic conditions, elevated temperature, drought, or the presence of light.

"Tissue-specific" promoters are transcriptional control elements that are only active in particular cells or tissues or organs, e.g., in plants or animals. Tissue-specific regulation may be achieved by certain intrinsic factors that ensure that genes encoding

proteins specific to a given tissue are expressed. Such factors are known to exist in mammals and plants so as to allow for specific tissues to develop.

The term "plant" includes whole plants, plant parts (e.g., leaves, stems, flowers, roots, etc.), plant protoplasts, seeds and plant cells and progeny of same. The class of plants which can be used in the method of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including angiosperms (monocotyledonous and dicotyledonous plants), as well as gymnosperms. It includes plants of a variety of ploidy levels, including polyploid, diploid, haploid and hemizygous states. As used herein, the term "transgenic plant" includes plants or plant cells into which a heterologous nucleic acid sequence has been inserted, e.g., the nucleic acids and various recombinant constructs (e.g., expression cassettes) of the invention.

"Plasmids" can be commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. Equivalent plasmids to those described herein are known in the art and will be apparent to the ordinarily skilled artisan.

"Amino acid" or "amino acid sequence" as used herein refer to an oligopeptide, peptide, polypeptide, or protein sequence, or to a fragment, portion, or subunit of any of these and to naturally occurring or synthetic molecules.

"Amino acid" or "amino acid sequence" include an oligopeptide, peptide, polypeptide, or protein sequence, or to a fragment, portion, or subunit of any of these, and to naturally occurring or synthetic molecules. The term "polypeptide" as used herein, refers to amino acids joined to each other by peptide bonds or modified peptide bonds, *i.e.*, peptide isosteres and may contain modified amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as post-translational processing, or by chemical modification techniques that are well known in the art.

Modifications can occur anywhere in the polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also a given polypeptide may have many types of modifications.

Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of a phosphatidylinositol, cross-linking cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of

pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, xylan hydrolase processing, phosphorylation, prenylation, racemization, selenoylation, sulfation and transfer-RNA mediated addition of amino acids to protein such as arginylation. (See
5 Creighton, T.E., Proteins – Structure and Molecular Properties 2nd Ed., W.H. Freeman and Company, New York (1993); *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York, pp. 1-12 (1983)). The peptides and polypeptides of the invention also include all “mimetic” and “peptidomimetic” forms, as described in further detail, below.

10 As used herein, the term “isolated” means that the material is removed from its original environment (*e.g.*, the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector
15 and/or such polynucleotides or polypeptides could be part of a composition and still be isolated in that such vector or composition is not part of its natural environment. As used herein, the term “purified” does not require absolute purity; rather, it is intended as a relative definition. Individual nucleic acids obtained from a library have been conventionally purified to electrophoretic homogeneity. The sequences obtained from these clones could not be obtained
20 directly either from the library or from total human DNA. The purified nucleic acids of the invention have been purified from the remainder of the genomic DNA in the organism by at least 10^4 - 10^6 fold. However, the term “purified” also includes nucleic acids that have been purified from the remainder of the genomic DNA or from other sequences in a library or other environment by at least one order of magnitude, typically two or three orders and more typically
25 four or five orders of magnitude.

As used herein, the term “recombinant” means that the nucleic acid is adjacent to a “backbone” nucleic acid to which it is not adjacent in its natural environment. Additionally, to be “enriched” the nucleic acids will represent 5% or more of the number of nucleic acid inserts in a population of nucleic acid backbone molecules. Backbone molecules according to the
30 invention include nucleic acids such as expression vectors, self-replicating nucleic acids, viruses, integrating nucleic acids and other vectors or nucleic acids used to maintain or manipulate a nucleic acid insert of interest. Typically, the enriched nucleic acids represent 15% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules. More typically, the enriched nucleic acids represent 50% or more of the number of nucleic acid inserts

in the population of recombinant backbone molecules. In a one aspect, the enriched nucleic acids represent 90% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules.

“Recombinant” polypeptides or proteins refer to polypeptides or proteins
5 produced by recombinant DNA techniques; *i.e.*, produced from cells transformed by an exogenous DNA construct encoding the desired polypeptide or protein. “Synthetic” polypeptides or protein are those prepared by chemical synthesis. Solid-phase chemical peptide synthesis methods can also be used to synthesize the polypeptide or fragments of the invention. Such method have been known in the art since the early 1960's (Merrifield, R. B., *J.*
10 *Am. Chem. Soc.*, 85:2149-2154, 1963) (See also Stewart, J. M. and Young, J. D., Solid Phase Peptide Synthesis, 2nd Ed., Pierce Chemical Co., Rockford, Ill., pp. 11-12)) and have recently been employed in commercially available laboratory peptide design and synthesis kits (Cambridge Research Biochemicals). Such commercially available laboratory kits have generally utilized the teachings of H. M. Geysen *et al*, *Proc. Natl. Acad. Sci., USA*, 81:3998
15 (1984) and provide for synthesizing peptides upon the tips of a multitude of “rods” or “pins” all of which are connected to a single plate. When such a system is utilized, a plate of rods or pins is inverted and inserted into a second plate of corresponding wells or reservoirs, which contain solutions for attaching or anchoring an appropriate amino acid to the pin's or rod's tips. By repeating such a process step, *i.e.*, inverting and inserting the rod's and pin's tips into appropriate
20 solutions, amino acids are built into desired peptides. In addition, a number of available Fmoc peptide synthesis systems are available. For example, assembly of a polypeptide or fragment can be carried out on a solid support using an Applied Biosystems, Inc. Model 431A automated peptide synthesizer. Such equipment provides ready access to the peptides of the invention, either by direct synthesis or by synthesis of a series of fragments that can be coupled using
25 other known techniques.

A promoter sequence is “operably linked to” a coding sequence when RNA polymerase which initiates transcription at the promoter will transcribe the coding sequence into mRNA.

“Plasmids” are designated by a lower case “p” preceded and/or followed by
30 capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described herein are known in the art and will be apparent to the ordinarily skilled artisan.

“Digestion” of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 µg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 µl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 µg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion, gel electrophoresis may be performed to isolate the desired fragment.

The phrase “substantially identical” in the context of two nucleic acids or polypeptides, refers to two or more sequences that have, e.g., at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more nucleotide or amino acid residue (sequence) identity, when compared and aligned for maximum correspondence, as measured using one of the known sequence comparison algorithms or by visual inspection. Typically, the substantial identity exists over a region of at least about 100 residues and most commonly the sequences are substantially identical over at least about 150-200 residues. In some aspects, the sequences are substantially identical over the entire length of the coding regions.

Additionally a “substantially identical” amino acid sequence is a sequence that differs from a reference sequence by one or more conservative or non-conservative amino acid substitutions, deletions, or insertions, particularly when such a substitution occurs at a site that is not the active site of the molecule and provided that the polypeptide essentially retains its functional properties. A conservative amino acid substitution, for example, substitutes one amino acid for another of the same class (e.g., substitution of one hydrophobic amino acid, such as isoleucine, valine, leucine, or methionine, for another, or substitution of one polar amino acid for another, such as substitution of arginine for lysine, glutamic acid for aspartic acid or glutamine for asparagine). One or more amino acids can be deleted, for example, from a xylanase polypeptide, resulting in modification of the structure of the polypeptide, without significantly altering its biological activity. For example, amino- or

carboxyl-terminal amino acids that are not required for xylanase biological activity can be removed. Modified polypeptide sequences of the invention can be assayed for xylanase biological activity by any number of methods, including contacting the modified polypeptide sequence with a xylanase substrate and determining whether the modified polypeptide
5 decreases the amount of specific substrate in the assay or increases the bioproducts of the enzymatic reaction of a functional xylanase polypeptide with the substrate.

“Fragments” as used herein are a portion of a naturally occurring protein which can exist in at least two different conformations. Fragments can have the same or substantially the same amino acid sequence as the naturally occurring protein. “Substantially
10 the same” means that an amino acid sequence is largely, but not entirely, the same, but retains at least one functional activity of the sequence to which it is related. In general two amino acid sequences are “substantially the same” or “substantially homologous” if they are at least about 85% identical. Fragments which have different three dimensional structures as the naturally occurring protein are also included. An example of this, is a “pro-form” molecule,
15 such as a low activity proprotein that can be modified by cleavage to produce a mature enzyme with significantly higher activity.

“Hybridization” refers to the process by which a nucleic acid strand joins with a complementary strand through base pairing. Hybridization reactions can be sensitive and selective so that a particular sequence of interest can be identified even in samples in which it
20 is present at low concentrations. Suitably stringent conditions can be defined by, for example, the concentrations of salt or formamide in the prehybridization and hybridization solutions, or by the hybridization temperature and are well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature. In alternative aspects,
25 nucleic acids of the invention are defined by their ability to hybridize under various stringency conditions (e.g., high, medium, and low), as set forth herein.

For example, hybridization under high stringency conditions could occur in about 50% formamide at about 37°C to 42°C. Hybridization could occur under reduced stringency conditions in about 35% to 25% formamide at about 30°C to 35°C. In particular,
30 hybridization could occur under high stringency conditions at 42°C in 50% formamide, 5X SSPE, 0.3% SDS and 200 n/ml sheared and denatured salmon sperm DNA. Hybridization could occur under reduced stringency conditions as described above, but in 35% formamide at a reduced temperature of 35°C. The temperature range corresponding to a particular level of stringency can be further narrowed by calculating the purine to pyrimidine ratio of the

nucleic acid of interest and adjusting the temperature accordingly. Variations on the above ranges and conditions are well known in the art.

The term "variant" refers to polynucleotides or polypeptides of the invention modified at one or more base pairs, codons, introns, exons, or amino acid residues (respectively) yet still retain the biological activity of a xylanase of the invention. Variants can be produced by any number of means included methods such as, for example, error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly (e.g., GeneReassembly™, see, e.g., U.S. Patent No. 6,537,776), GSSM™ and any combination thereof.

Table 1 and Table 2 list variants obtained by mutating SEQ ID NO:189 (encoding SEQ ID NO:190) by GSSM™. The invention provides nucleic acids having one or more, or all, of the sequences as set forth in Tables 1 and 2, i.e., nucleic acids having sequences that are variants of SEQ ID NO:189, where the variations are set forth in Table 1 and Table 2, and the polypeptides that are encoded by these variants.

These GSSM™ variants (set forth in Tables 1 and 2) were tested for thermal tolerance (see Examples, below). Mutants D, F, G, H, I, J, K, S, T, U, V, W, X, Y, Z, AA, DD and EE were found to have the highest thermal tolerance among the mutants in Table 1. Mutants may also be combined to form a larger mutant. For example, mutants D, F, H, I, S, V, X and AA of Table 1 were combined to form a larger mutant termed "8x" with a sequence as set forth in SEQ ID NO:375 (polypeptide encoding nucleic acid) and SEQ ID NO:376 (amino acid sequence). Figure 5 is a graph comparing the activity of the wild type sequence (SEQ ID NOS: 189 and 190) to the 8x mutant (SEQ ID NOS: 259 and 260). In comparing the wild type and the 8x mutant, it was discovered that the optimal temperature for both was 65°C and that the optimal pH for both was 5.5. The wild type sequence was found to maintain its stability for less than 1 minute at 65°C, while the 8x mutant (SEQ ID NOS:375, 376) was found to maintain its stability for more than 10 minutes at 85°C. The substrate used was AZO-AZO-xylan. In one aspect, the 8x mutant (SEQ ID NOS:375, 376) was evolved by GSSM™. In another aspect, the wild type is a GSSM™ parent for thermal tolerance evolution.

Table 1

Mutant	Mutation	Wild type Seq	GSSM™ Seq
A	A2F	GCC	TTT
B	A2D	GCC	GAC
C	A5H	GCT	CAC
D	D8F	GAC	TTC
E	Q11L	CAA	CTC
F	Q11H	CAA	CAC
G	N12L	AAT	TTG
H	N12L	AAT	TTG
I	G17I	GGT	ATA
J	Q11H,T23T	CAA,ACC	CAT,ACG
K	Q11H	CAA	CAT
L	S26P	TCT	CCG
M	S26P	TCT	CCA
N	S35F	TCA	TTT
O	No Change	GTT	GTA
P	A51P	GCA	CCG
Q	A51P	GCA	CCG
R	G60R	GGA	CGC
S	G60H	GGA	CAC
T	G60H	GGA	CAC
U	P64C	CCG	TGT
V	P64V	CCG	GTA
W	P64V	CCG	GTT
X	S65V	TCC	GTG
Y	Q11H	CAA	CAT
Z	G68I	GGA	ATA
AA	G68A	GGA	GCT
BB	A71G	GCT	GGA
CC	No Change	AAT	AAC
DD	S79P	TCA	CCA
EE	S79P	TCA	CCC
FF	T95S	ACT	TCT
GG	Y98P	TAT	CCG
HH	T114S	ACT	AGC
II	No Change	AAC	AAC
JJ	No Change	AGG	AGA
KK	I142L	ATT	CTG
LL	S151I	AGC	ATC
MM	S138T,S151A	TCG,AGC	ACG,GCG
NN	K158R	AAG	CGG
OO	K160V,V172I	AAA,GTA	GTT,ATA

The codon variants as set forth in Table 2 that produced variants (of SEQ ID NO:189) with the best variation or “improvement” over “wild type” (SEQ ID NO:189) in thermal tolerance are highlighted. As noted above, the invention provides nucleic acids, and the polypeptides that encode them, comprising one, several or all or the variations set forth in Table 2 and Table 1.

Table 2

<u>Mutation</u>	<u>Wild type Sequence</u>	<u>GSSM™ Sequence</u>	<u>Other codons also coding for same changed amino acid</u>
A2F	GCC	TTT	TTC
A2D	GCC	GAC	GAT
A5H	GCT	CAC	CAT
D3F	GAC	TTT	TTT
Q11L	CAA	CTC	TTA, TTG, CTT, CTA, CTG
Q11E	CAA	CAC, CAT	-
N12I	AAT	TTG	TTA, CTC, CTT, CTA, CTG
G17	GGT	ATA	ATT, ATC
T23T	ACC	ACG	ACT, ACC, ACA
S26P	TCT	CCG, CCA	CCC
S35F	TCA	TTT	TTC
A51P	GCA	CCG	CCC, CCA
G60R	GGA	CGC	CGT, CGA, CGG, AGA, AGG
E60H	GGA	CAC	CAT
E61C	CCG	TCG	TGC
E64V	CCG	GTA, GTT	GTC, GTG
S65V	TCC	GTC	GTC, GTA, GTT
E68I	GGA	ATA	ATT, ATC
E68A	GGA	GCA	GCG, GCC, GCA
A71G	GCT	GGA	GGT, GGC, GGG
S72I	TCA	TGA, TCG	CCG
T95S	ACT	TCT	TCC, TCA, TCG, AGT, AGC
Y98P	TAT	CCG	CCC, CCA
T114S	ACT	AGC	TCC, TCA, TCG, AGT, TCT
I142L	ATT	CTG	TTA, CTC, CTT, CTA, TTG
S151I	AGC	ATC	ATT, ATA
S138T	TCG	ACG	ACT, ACC, ACA
S151A	AGC	GCG	GCT, GCC, GCA
K158R	AAG	CGG	CGT, CGA, CGC, AGA, AGG
K160V	AAA	GTT	GTC, GTA, GTG
V172I	GTA	ATA	ATT, ATC

In one aspect the amino acid sequence of an amino acid sequence (SEQ ID NO: 208) of Group B amino acid sequences is modified by a single amino acid mutation. In a specific aspect, that mutation is an asparagine to aspartic acid mutation. The resulting amino acid sequence and corresponding nucleic acid sequence are set forth as SEQ ID NO:252 and SEQ ID NO:251, respectively. Single amino acid mutations with an improvement in the pH optimum of the enzyme, such as the mutation of SEQ ID NO:208, have been shown in the art with respect to xylanases. (See, for example, Joshi, M., Sidhu, G., Pot, I., Brayer, G., Withers, S., McIntosh, L., *J. Mol. Bio.* 299, 255-279 (2000).) It is also noted that in such single amino acid mutations, portions of the sequences may be removed in the subcloning process. For example, SEQ ID NO:207 and SEQ ID NO:251 differ in only

one nucleotide, over the area that the sequences align. However, it is noted that a 78 nucleotide area at the N-terminus of SEQ ID NO:207 was removed from the N-terminus of SEQ ID NO:251 in the subcloning. Additionally, the first three nucleotides in SEQ ID NO:251 were changed to ATG and then the point mutation was made at the sixth nucleotide in SEQ ID NO:251.

The term "saturation mutagenesis", "gene site saturated mutagenesis" or "GSSM™" includes a method that uses degenerate oligonucleotide primers to introduce point mutations into a polynucleotide, as described in detail, below.

The term "optimized directed evolution system" or "optimized directed evolution" includes a method for reassembling fragments of related nucleic acid sequences, e.g., related genes, and explained in detail, below.

The term "synthetic ligation reassembly" or "SLR" includes a method of ligating oligonucleotide fragments in a non-stochastic fashion, and explained in detail, below.

Generating and Manipulating Nucleic Acids

The invention provides nucleic acids (e.g., SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193,

SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203,
SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213,
SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223,
SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233,
5 SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243,
SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253,
SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263,
SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273,
SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283,
10 SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293,
SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303,
SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313,
SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323,
SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333,
15 SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343,
SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353,
SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363,
SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373,
SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379; nucleic acids encoding polypeptides
20 as set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10,
SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID
NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32,
SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID
NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54,
25 SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID
NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76,
SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID
NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98,
SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108,
30 SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118,
SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128,
SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138;
SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, NO:146, SEQ ID NO:148, SEQ ID
NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID

NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID
NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID
NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188, SEQ ID
NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198, SEQ ID
5 NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID
NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218, SEQ ID
NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228, SEQ ID
NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID
NO:240, SEQ ID NO:242, SEQ ID NO:244, SEQ ID NO:246, SEQ ID NO:248, SEQ ID
10 NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256, SEQ ID NO:258, SEQ ID
NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:268, SEQ ID
NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276, SEQ ID NO:278, SEQ ID
NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286, SEQ ID NO:288, SEQ ID
NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296, SEQ ID NO:298, SEQ ID
15 NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306, SEQ ID NO:308, SEQ ID
NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316, SEQ ID NO:318, SEQ ID
NO:320, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328, SEQ ID
NO:330, SEQ ID NO:332, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:338, SEQ ID
NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348, SEQ ID
20 NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358, SEQ ID
NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368, SEQ ID
NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378 or SEQ ID
NO:380), including expression cassettes such as expression vectors, encoding the

polypeptides of the invention. The invention also includes methods for discovering new
25 xylanase sequences using the nucleic acids of the invention. The invention also includes
methods for inhibiting the expression of xylanase genes, transcripts and polypeptides using
the nucleic acids of the invention. Also provided are methods for modifying the nucleic acids
of the invention by, e.g., synthetic ligation reassembly, optimized directed evolution system
and/or saturation mutagenesis.

30 The nucleic acids of the invention can be made, isolated and/or manipulated
by, e.g., cloning and expression of cDNA libraries, amplification of message or genomic
DNA by PCR, and the like. For example, the following exemplary sequences of the
invention were initially derived from the following sources:

Table 3

	<u>SEQ ID</u>	<u>SOURCE</u>
	1, 2	Bacteria
	101, 102	Environmental
5	103, 104	Bacteria
	105, 106	Environmental
	107, 108	Bacteria
	109, 110	Environmental
	11, 12	Environmental
10	111, 112	Environmental
	113, 114	Environmental
	115, 116	Environmental
	117, 118	Environmental
	119, 120	Environmental
15	121, 122	Environmental
	123, 124	Environmental
	125, 126	Environmental
	127, 128	Environmental
	129, 130	Bacteria
20	13, 14	Environmental
	131, 132	Environmental
	133, 134	Environmental
	135, 136	Environmental
	137, 138	Environmental
25	139, 140	Environmental
	141, 142	Environmental
	143, 144	Bacteria
	145, 146	Eukaryote
	147, 148	Environmental
30	149, 150	Environmental
	15, 16	Environmental
	151, 152	Environmental
	153, 154	Environmental
	155, 156	Environmental
35	157, 158	Environmental
	159, 160	Environmental
	161, 162	Environmental
	163, 164	Environmental
	165, 166	Environmental
40	167, 168	Environmental
	169, 170	Environmental
	17, 18	Bacteria
	171, 172	Environmental
	173, 174	Environmental
45	175, 176	Environmental
	177, 178	Environmental
	179, 180	Environmental
	181, 182	Environmental
	183, 184	Environmental
50	185, 186	Environmental

	187, 188	Environmental
	189, 190	Environmental
	19, 20	Environmental
5	191, 192	Environmental
	193, 194	Environmental
	195, 196	Environmental
	197, 198	Environmental
	199, 200	Environmental
10	201, 202	Environmental
	203, 204	Environmental
	205, 206	Environmental
	207, 208	Environmental
	209, 210	Environmental
	21, 22	Environmental
15	211, 212	Environmental
	213, 214	Environmental
	215, 216	Environmental
	217, 218	Environmental
	219, 220	Environmental
20	221, 222	Environmental
	223, 224	Environmental
	225, 226	Environmental
	227, 228	Environmental
	229, 230	Environmental
25	23, 24	Environmental
	231, 232	Bacteria
	233, 234	Environmental
	235, 236	Environmental
	237, 238	Environmental
30	239, 240	Environmental
	241, 242	Environmental
	243, 244	Environmental
	245, 246	Environmental
	247, 248	Environmental
35	249, 250	Environmental
	25, 26	Environmental
	251, 252	Environmental
	253, 254	Environmental
	255, 256	Environmental
40	257, 258	Environmental
	259, 260	Environmental
	261, 262	Environmental
	263, 264	Environmental
	265, 266	Environmental
45	267, 268	Bacteria
	269, 270	Environmental
	27, 28	Environmental
	271, 272	Environmental
	273, 274	Environmental
50	275, 276	Environmental

	277, 278	Environmental
	279, 280	Environmental
	281, 282	Environmental
	283, 284	Environmental
5	285, 286	Environmental
	287, 288	Environmental
	289, 290	Environmental
	29, 30	Archaea
	291, 292	Environmental
10	293, 294	Environmental
	295, 296	Environmental
	297, 298	Environmental
	299, 300	Environmental
	3, 4	Environmental
15	301, 302	Environmental
	303, 304	Environmental
	305, 306	Bacteria
	307, 308	Environmental
	309, 310	Environmental
20	31, 32	Environmental
	311, 312	Environmental
	313, 314	Bacteria
	315, 316	Environmental
	317, 318	Environmental
25	319, 320	Environmental
	321, 322	Environmental
	323, 324	Environmental
	325, 326	Environmental
	327, 328	Environmental
30	329, 330	Environmental
	33, 34	Environmental
	331, 332	Environmental
	333, 334	Environmental
	335, 336	Environmental
35	337, 338	Environmental
	339, 340	Environmental
	341, 342	Environmental
	343, 344	Environmental
	345, 346	Environmental
40	347, 348	Environmental
	349, 350	Environmental
	35, 36	Environmental
	351, 352	Environmental
	353, 354	Environmental
45	355, 356	Environmental
	357, 358	Environmental
	359, 360	Environmental
	361, 362	Environmental
	363, 364	Environmental
50	365, 366	Environmental

	367, 368	Environmental
	369, 370	Environmental
	37, 38	Environmental
5	371, 372	Environmental
	373, 374	Environmental
	375, 376	Artificial
	377, 378	Artificial
	39, 40	Environmental
	41, 42	Environmental
10	43, 44	Environmental
	45, 46	Environmental
	47, 48	Environmental
	49, 50	Environmental
	5, 6	Environmental
15	51, 52	Environmental
	53, 54	Bacteria
	55, 56	Environmental
	57, 58	Environmental
	59, 60	Environmental
20	61, 62	Environmental
	63, 64	Environmental
	65, 66	Environmental
	67, 68	Environmental
	69, 70	Environmental
25	7, 8	Environmental
	71, 72	Environmental
	73, 74	Environmental
	75, 76	Environmental
	77, 78	Environmental
30	79, 80	Environmental
	81, 82	Environmental
	83, 84	Environmental
	85, 86	Bacteria
	87, 88	Environmental
35	89, 90	Bacteria
	9, 10	Environmental
	91, 92	Environmental
	93, 94	Environmental
	95, 96	Environmental
40	97, 98	Environmental
	99, 100	Environmental

In one aspect, the invention also provides xylanase-encoding nucleic acids with a common novelty in that they are derived from an environmental source, or a bacterial source, or an archaeal source.

45 In practicing the methods of the invention, homologous genes can be modified by manipulating a template nucleic acid, as described herein. The invention can be practiced

in conjunction with any method or protocol or device known in the art, which are well described in the scientific and patent literature.

One aspect of the invention is an isolated nucleic acid comprising one of the sequences of Group A nucleic acid sequences and sequences substantially identical thereto, the sequences complementary thereto, or a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive bases of one of the sequences of a Group A nucleic acid sequence (or the sequences complementary thereto). The isolated, nucleic acids may comprise DNA, including cDNA, genomic DNA and synthetic DNA. The DNA may be double-stranded or single-stranded and if single stranded may be the coding strand or non-coding (anti-sense) strand. Alternatively, the isolated nucleic acids may comprise RNA.

As discussed in more detail below, the isolated nucleic acids of one of the Group A nucleic acid sequences and sequences substantially identical thereto, may be used to prepare one of the polypeptides of a Group B amino acid sequence and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto.

Accordingly, another aspect of the invention is an isolated nucleic acid which encodes one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of one of the polypeptides of the Group B amino acid sequences. The coding sequences of these nucleic acids may be identical to one of the coding sequences of one of the nucleic acids of Group A nucleic acid sequences, or a fragment thereof or may be different coding sequences which encode one of the polypeptides of Group B amino acid sequences, sequences substantially identical thereto and fragments having at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids of one of the polypeptides of Group B amino acid sequences, as a result of the redundancy or degeneracy of the genetic code. The genetic code is well known to those of skill in the art and can be obtained, for example, on page 214 of B. Lewin, Genes VI, Oxford University Press, 1997.

The isolated nucleic acid which encodes one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, may include, but is not limited to: only the coding sequence of one of Group A nucleic acid sequences and sequences substantially identical thereto and additional coding sequences, such as leader sequences or

proprotein sequences and non-coding sequences, such as introns or non-coding sequences 5' and/or 3' of the coding sequence. Thus, as used herein, the term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only the coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

Alternatively, the nucleic acid sequences of Group A nucleic acid sequences and sequences substantially identical thereto, may be mutagenized using conventional techniques, such as site directed mutagenesis, or other techniques familiar to those skilled in the art, to introduce silent changes into the polynucleotides of Group A nucleic acid sequences and sequences substantially identical thereto. As used herein, "silent changes" include, for example, changes which do not alter the amino acid sequence encoded by the polynucleotide. Such changes may be desirable in order to increase the level of the polypeptide produced by host cells containing a vector encoding the polypeptide by introducing codons or codon pairs which occur frequently in the host organism.

The invention also relates to polynucleotides which have nucleotide changes which result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptides of Group B amino acid sequences and sequences substantially identical thereto. Such nucleotide changes may be introduced using techniques such as site directed mutagenesis, random chemical mutagenesis, exonuclease III deletion and other recombinant DNA techniques. Alternatively, such nucleotide changes may be naturally occurring allelic variants which are isolated by identifying nucleic acids which specifically hybridize to probes comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive bases of one of the sequences of Group A nucleic acid sequences and sequences substantially identical thereto (or the sequences complementary thereto) under conditions of high, moderate, or low stringency as provided herein.

General Techniques

The nucleic acids used to practice this invention, whether RNA, iRNA, antisense nucleic acid, cDNA, genomic DNA, vectors, viruses or hybrids thereof, may be isolated from a variety of sources, genetically engineered, amplified, and/or expressed/generated recombinantly. Recombinant polypeptides (e.g., xylanases) generated from these nucleic acids can be individually isolated or cloned and tested for a desired activity. Any recombinant expression system can be used, including bacterial, mammalian, yeast, insect or plant cell expression systems.

Alternatively, these nucleic acids can be synthesized *in vitro* by well-known chemical synthesis techniques, as described in, e.g., Adams (1983) J. Am. Chem. Soc. 105:661; Belousov (1997) Nucleic Acids Res. 25:3440-3444; Frenkel (1995) Free Radic. Biol. Med. 19:373-380; Blommers (1994) Biochemistry 33:7886-7896; Narang (1979) Meth. Enzymol. 68:90; Brown (1979) Meth. Enzymol. 68:109; Beaucage (1981) Tetra. Lett. 22:1859; U.S. Patent No. 4,458,066.

Techniques for the manipulation of nucleic acids, such as, e.g., subcloning, labeling probes (e.g., random-primer labeling using Klenow polymerase, nick translation, amplification), sequencing, hybridization and the like are well described in the scientific and patent literature, see, e.g., Sambrook, ed., MOLECULAR CLONING: A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Ausubel, ed. John Wiley & Sons, Inc., New York (1997); LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY: HYBRIDIZATION WITH NUCLEIC ACID PROBES, Part I. Theory and Nucleic Acid Preparation, Tijssen, ed. Elsevier, N.Y. (1993).

Another useful means of obtaining and manipulating nucleic acids used to practice the methods of the invention is to clone from genomic samples, and, if desired, screen and re-clone inserts isolated or amplified from, e.g., genomic clones or cDNA clones. Sources of nucleic acid used in the methods of the invention include genomic or cDNA libraries contained in, e.g., mammalian artificial chromosomes (MACs), see, e.g., U.S. Patent Nos. 5,721,118; 6,025,155; human artificial chromosomes, see, e.g., Rosenfeld (1997) Nat. Genet. 15:333-335; yeast artificial chromosomes (YAC); bacterial artificial chromosomes (BAC); P1 artificial chromosomes, see, e.g., Woon (1998) Genomics 50:306-316; P1-derived vectors (PACs), see, e.g., Kern (1997) Biotechniques 23:120-124; cosmids, recombinant viruses, phages or plasmids.

In one aspect, a nucleic acid encoding a polypeptide of the invention is assembled in appropriate phase with a leader sequence capable of directing secretion of the translated polypeptide or fragment thereof.

The invention provides fusion proteins and nucleic acids encoding them. A polypeptide of the invention can be fused to a heterologous peptide or polypeptide, such as N-terminal identification peptides which impart desired characteristics, such as increased stability or simplified purification. Peptides and polypeptides of the invention can also be synthesized and expressed as fusion proteins with one or more additional domains linked thereto for, e.g., producing a more immunogenic peptide, to more readily isolate a

recombinantly synthesized peptide, to identify and isolate antibodies and antibody-expressing B cells, and the like. Detection and purification facilitating domains include, e.g., metal chelating peptides such as polyhistidine tracts and histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on
5 immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequences such as Factor Xa or enterokinase (Invitrogen, San Diego CA) between a purification domain and the motif-comprising peptide or polypeptide to facilitate purification. For example, an expression vector can include an epitope-encoding nucleic acid sequence
10 linked to six histidine residues followed by a thioredoxin and an enterokinase cleavage site (see e.g., Williams (1995) Biochemistry 34:1787-1797; Dobeli (1998) Protein Expr. Purif. 12:404-414). The histidine residues facilitate detection and purification while the enterokinase cleavage site provides a means for purifying the epitope from the remainder of the fusion protein. Technology pertaining to vectors encoding fusion proteins and application
15 of fusion proteins are well described in the scientific and patent literature, see e.g., Kroll (1993) DNA Cell. Biol., 12:441-53.

Transcriptional and translational control sequences

The invention provides nucleic acid (e.g., DNA) sequences of the invention operatively linked to expression (e.g., transcriptional or translational) control sequence(s),
20 e.g., promoters or enhancers, to direct or modulate RNA synthesis/ expression. The expression control sequence can be in an expression vector. Exemplary bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, PL and trp. Exemplary eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein I.

Promoters suitable for expressing a polypeptide in bacteria include the *E. coli* lac or trp promoters, the lacI promoter, the lacZ promoter, the T3 promoter, the T7 promoter, the gpt promoter, the lambda PR promoter, the lambda PL promoter, promoters from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), and the acid phosphatase promoter. Eukaryotic promoters include the CMV immediate early promoter,
30 the HSV thymidine kinase promoter, heat shock promoters, the early and late SV40 promoter, LTRs from retroviruses, and the mouse metallothionein-I promoter. Other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses may also be used. Promoters suitable for expressing the polypeptide or fragment thereof in bacteria

include the *E. coli lac* or *trp* promoters, the *lacI* promoter, the *lacZ* promoter, the *T3* promoter, the *T7* promoter, the *gpt* promoter, the *lambda P_R* promoter, the *lambda P_L* promoter, promoters from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK) and the acid phosphatase promoter. Fungal promoters include the α factor promoter. Eukaryotic promoters include the CMV immediate early promoter, the HSV thymidine kinase promoter, heat shock promoters, the early and late SV40 promoter, LTRs from retroviruses and the mouse metallothionein-I promoter. Other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses may also be used.

10 *Tissue-Specific Plant Promoters*

The invention provides expression cassettes that can be expressed in a tissue-specific manner, e.g., that can express a xylanase of the invention in a tissue-specific manner. The invention also provides plants or seeds that express a xylanase of the invention in a tissue-specific manner. The tissue-specificity can be seed specific, stem specific, leaf
15 specific, root specific, fruit specific and the like.

In one aspect, a constitutive promoter such as the CaMV 35S promoter can be used for expression in specific parts of the plant or seed or throughout the plant. For example, for overexpression, a plant promoter fragment can be employed which will direct expression of a nucleic acid in some or all tissues of a plant, e.g., a regenerated plant. Such
20 promoters are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, and other transcription initiation regions from various plant genes known to those of skill.

Such genes include, e.g., *ACT11* from *Arabidopsis* (Huang (1996) *Plant Mol. Biol.* 33:125-139); *Cat3* from *Arabidopsis* (GenBank No. U43147, Zhong (1996) *Mol. Gen. Genet.* 251:196-203); the gene encoding stearyl-acyl carrier protein desaturase from *Brassica napus* (Genbank No. X74782, Solocombe (1994) *Plant Physiol.* 104:1167-1176); *GPc1* from maize (GenBank No. X15596; Martinez (1989) *J. Mol. Biol.* 208:551-565); the *Gpc2* from maize
25 (GenBank No. U45855, Manjunath (1997) *Plant Mol. Biol.* 33:97-112); plant promoters described in U.S. Patent Nos. 4,962,028; 5,633,440.

The invention uses tissue-specific or constitutive promoters derived from viruses which can include, e.g., the tobamovirus subgenomic promoter (Kumagai (1995)

Proc. Natl. Acad. Sci. USA 92:1679-1683; the rice tungro bacilliform virus (RTBV), which replicates only in phloem cells in infected rice plants, with its promoter which drives strong phloem-specific reporter gene expression; the cassava vein mosaic virus (CVMV) promoter, with highest activity in vascular elements, in leaf mesophyll cells, and in root tips (Verdaguer
5 (1996) Plant Mol. Biol. 31:1129-1139).

Alternatively, the plant promoter may direct expression of xylanase-expressing nucleic acid in a specific tissue, organ or cell type (*i.e.* tissue-specific promoters) or may be otherwise under more precise environmental or developmental control or under the control of an inducible promoter. Examples of environmental conditions that may affect
10 transcription include anaerobic conditions, elevated temperature, the presence of light, or sprayed with chemicals/hormones. For example, the invention incorporates the drought-inducible promoter of maize (Busk (1997) *supra*); the cold, drought, and high salt inducible promoter from potato (Kirch (1997) Plant Mol. Biol. 33:897 909).

Tissue-specific promoters can promote transcription only within a certain time
15 frame of developmental stage within that tissue. See, e.g., Blazquez (1998) Plant Cell 10:791-800, characterizing the *Arabidopsis* LEAFY gene promoter. See also Cardon (1997) Plant J 12:367-77, describing the transcription factor SPL3, which recognizes a conserved sequence motif in the promoter region of the *A. thaliana* floral meristem identity gene AP1; and Mandel (1995) Plant Molecular Biology, Vol. 29, pp 995-1004, describing the meristem
20 promoter eIF4. Tissue specific promoters which are active throughout the life cycle of a particular tissue can be used. In one aspect, the nucleic acids of the invention are operably linked to a promoter active primarily only in cotton fiber cells. In one aspect, the nucleic acids of the invention are operably linked to a promoter active primarily during the stages of cotton fiber cell elongation, e.g., as described by Rinehart (1996) *supra*. The nucleic acids
25 can be operably linked to the Fb12A gene promoter to be preferentially expressed in cotton fiber cells (Ibid) . See also, John (1997) Proc. Natl. Acad. Sci. USA 89:5769-5773; John, et al., U.S. Patent Nos. 5,608,148 and 5,602,321, describing cotton fiber-specific promoters and methods for the construction of transgenic cotton plants. Root-specific promoters may also be used to express the nucleic acids of the invention. Examples of root-specific promoters
30 include the promoter from the alcohol dehydrogenase gene (DeLisle (1990) Int. Rev. Cytol. 123:39-60). Other promoters that can be used to express the nucleic acids of the invention include, e.g., ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed coat-specific promoters, or some combination thereof; a leaf-specific promoter (see, e.g., Busk (1997) Plant J. 11:1285 1295, describing a leaf-specific promoter in maize); the ORF13

promoter from *Agrobacterium rhizogenes* (which exhibits high activity in roots, see, e.g., Hansen (1997) supra); a maize pollen specific promoter (see, e.g., Guerrero (1990) Mol. Gen. Genet. 224:161 168); a tomato promoter active during fruit ripening, senescence and abscission of leaves and, to a lesser extent, of flowers can be used (see, e.g., Blume (1997) Plant J. 12:731 746); a pistil-specific promoter from the potato SK2 gene (see, e.g., Ficker (1997) Plant Mol. Biol. 35:425 431); the Blec4 gene from pea, which is active in epidermal tissue of vegetative and floral shoot apices of transgenic alfalfa making it a useful tool to target the expression of foreign genes to the epidermal layer of actively growing shoots or fibers; the ovule-specific BEL1 gene (see, e.g., Reiser (1995) Cell 83:735-742, GenBank No. U39944); and/or, the promoter in Klee, U.S. Patent No. 5,589,583, describing a plant promoter region is capable of conferring high levels of transcription in meristematic tissue and/or rapidly dividing cells.

Alternatively, plant promoters which are inducible upon exposure to plant hormones, such as auxins, are used to express the nucleic acids of the invention. For example, the invention can use the auxin-response elements E1 promoter fragment (AuxREs) in the soybean (*Glycine max* L.) (Liu (1997) Plant Physiol. 115:397-407); the auxin-responsive *Arabidopsis* GST6 promoter (also responsive to salicylic acid and hydrogen peroxide) (Chen (1996) Plant J. 10: 955-966); the auxin-inducible parC promoter from tobacco (Sakai (1996) 37:906-913); a plant biotin response element (Streit (1997) Mol. Plant Microbe Interact. 10:933-937); and, the promoter responsive to the stress hormone abscisic acid (Sheen (1996) Science 274:1900-1902).

The nucleic acids of the invention can also be operably linked to plant promoters which are inducible upon exposure to chemicals reagents which can be applied to the plant, such as herbicides or antibiotics. For example, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, can be used (De Veylder (1997) Plant Cell Physiol. 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequence can be under the control of, e.g., a tetracycline-inducible promoter, e.g., as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) Plant J. 11:465-473); or, a salicylic acid-responsive element (Stange (1997) Plant J. 11:1315-1324). Using chemically- (e.g., hormone- or pesticide-) induced promoters, i.e., promoter responsive to a chemical which can be applied to the transgenic plant in the field, expression of a polypeptide of the invention can be induced at a particular stage of development of the plant. Thus, the invention also

provides for transgenic plants containing an inducible gene encoding for polypeptides of the invention whose host range is limited to target plant species, such as corn, rice, barley, wheat, potato or other crops, inducible at any stage of development of the crop.

One of skill will recognize that a tissue-specific plant promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, a tissue-specific promoter is one that drives expression preferentially in the target tissue or cell type, but may also lead to some expression in other tissues as well.

The nucleic acids of the invention can also be operably linked to plant promoters which are inducible upon exposure to chemicals reagents. These reagents include, e.g., herbicides, synthetic auxins, or antibiotics which can be applied, e.g., sprayed, onto transgenic plants. Inducible expression of the xylanase-producing nucleic acids of the invention will allow the grower to select plants with the optimal xylanase expression and/or activity. The development of plant parts can thus controlled. In this way the invention provides the means to facilitate the harvesting of plants and plant parts. For example, in various embodiments, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, is used (De Veylder (1997) Plant Cell Physiol. 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequences of the invention are also under the control of a tetracycline-inducible promoter, e.g., as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) Plant J. 11:465-473); or, a salicylic acid-responsive element (Stange (1997) Plant J. 11:1315-1324).

In some aspects, proper polypeptide expression may require polyadenylation region at the 3'-end of the coding region. The polyadenylation region can be derived from the natural gene, from a variety of other plant (or animal or other) genes, or from genes in the *Agrobacterial* T-DNA.

Expression vectors and cloning vehicles

The invention provides expression vectors and cloning vehicles comprising nucleic acids of the invention, e.g., sequences encoding the xylanases of the invention.

Expression vectors and cloning vehicles of the invention can comprise viral particles, baculovirus, phage, plasmids, phagemids, cosmids, fosmids, bacterial artificial chromosomes, viral DNA (e.g., vaccinia, adenovirus, fowl pox virus, pseudorabies and derivatives of SV40), P1-based artificial chromosomes, yeast plasmids, yeast artificial chromosomes, and any other

vectors specific for specific hosts of interest (such as bacillus, Aspergillus and yeast). Vectors of the invention can include chromosomal, non-chromosomal and synthetic DNA sequences. Large numbers of suitable vectors are known to those of skill in the art, and are commercially available. Exemplary vectors are include: bacterial: pQE vectors (Qiagen),
5 pBluescript plasmids, pNH vectors, (lambda-ZAP vectors (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eukaryotic: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). However, any other plasmid or other vector may be used so long as they are replicable and viable in the host. Low copy number or high copy number vectors may be employed with the present invention.

10 The expression vector can comprise a promoter, a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression. Mammalian expression vectors can comprise an origin of replication, any necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed
15 sequences. In some aspects, DNA sequences derived from the SV40 splice and polyadenylation sites may be used to provide the required non-transcribed genetic elements.

In one aspect, the expression vectors contain one or more selectable marker genes to permit selection of host cells containing the vector. Such selectable markers include genes encoding dihydrofolate reductase or genes conferring neomycin resistance for
20 eukaryotic cell culture, genes conferring tetracycline or ampicillin resistance in *E. coli*, and the *S. cerevisiae* TRP1 gene. Promoter regions can be selected from any desired gene using chloramphenicol transferase (CAT) vectors or other vectors with selectable markers.

Vectors for expressing the polypeptide or fragment thereof in eukaryotic cells can also contain enhancers to increase expression levels. Enhancers are cis-acting elements
25 of DNA, usually from about 10 to about 300 bp in length that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and the adenovirus enhancers.

A nucleic acid sequence can be inserted into a vector by a variety of
30 procedures. In general, the sequence is ligated to the desired position in the vector following digestion of the insert and the vector with appropriate restriction endonucleases. Alternatively, blunt ends in both the insert and the vector may be ligated. A variety of cloning techniques are known in the art, e.g., as described in Ausubel and Sambrook. Such procedures and others are deemed to be within the scope of those skilled in the art.

The vector can be in the form of a plasmid, a viral particle, or a phage. Other vectors include chromosomal, non-chromosomal and synthetic DNA sequences, derivatives of SV40; bacterial plasmids, phage DNA, baculovirus, yeast plasmids, vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. A variety of cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by, e.g., Sambrook.

Particular bacterial vectors which can be used include the commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden), GEM1 (Promega Biotec, Madison, WI, USA) pQE70, pQE60, pQE-9 (Qiagen), pD10, psiX174 pBluescript II KS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, DR540, pRIT5 (Pharmacia), pKK232-8 and pCM7. Particular eukaryotic vectors include pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, and pSVL (Pharmacia). However, any other vector may be used as long as it is replicable and viable in the host cell.

The nucleic acids of the invention can be expressed in expression cassettes, vectors or viruses and transiently or stably expressed in plant cells and seeds. One exemplary transient expression system uses episomal expression systems, e.g., cauliflower mosaic virus (CaMV) viral RNA generated in the nucleus by transcription of an episomal mini-chromosome containing supercoiled DNA, see, e.g., Covey (1990) Proc. Natl. Acad. Sci. USA 87:1633-1637. Alternatively, coding sequences, i.e., all or sub-fragments of sequences of the invention can be inserted into a plant host cell genome becoming an integral part of the host chromosomal DNA. Sense or antisense transcripts can be expressed in this manner. A vector comprising the sequences (e.g., promoters or coding regions) from nucleic acids of the invention can comprise a marker gene that confers a selectable phenotype on a plant cell or a seed. For example, the marker may encode biocide resistance, particularly antibiotic resistance, such as resistance to kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or Basta.

Expression vectors capable of expressing nucleic acids and proteins in plants are well known in the art, and can include, e.g., vectors from *Agrobacterium* spp., potato virus X (see, e.g., Angell (1997) EMBO J. 16:3675-3684), tobacco mosaic virus (see, e.g., Casper (1996) Gene 173:69-73), tomato bushy stunt virus (see, e.g., Hillman (1989) Virology 169:42-50), tobacco etch virus (see, e.g., Dolja (1997) Virology 234:243-252), bean golden mosaic virus (see, e.g., Morinaga (1993) Microbiol Immunol. 37:471-476), cauliflower mosaic virus (see, e.g., Cecchini (1997) Mol. Plant Microbe Interact. 10:1094-1101), maize

Ac/Ds transposable element (see, *e.g.*, Rubin (1997) Mol. Cell. Biol. 17:6294-6302; Kunze (1996) Curr. Top. Microbiol. Immunol. 204:161-194), and the maize suppressor-mutator (Spm) transposable element (see, *e.g.*, Schlappi (1996) Plant Mol. Biol. 32:717-725); and derivatives thereof.

5 In one aspect, the expression vector can have two replication systems to allow it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector can contain at least one sequence homologous to the host cell genome. It can contain two homologous sequences which flank the expression
10 construct. The integrating vector can be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating vectors are well known in the art.

Expression vectors of the invention may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed, *e.g.*, genes which
15 render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and tetracycline. Selectable markers can also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct RNA synthesis. Particular
20 named bacterial promoters include *lacI*, *lacZ*, *T3*, *T7*, *gpt*, *lambda P_R*, *P_L* and *trp*. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The
25 vector may also include appropriate sequences for amplifying expression. Promoter regions can be selected from any desired gene using chloramphenicol transferase (CAT) vectors or other vectors with selectable markers. In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture,
30 or such as tetracycline or ampicillin resistance in *E. coli*.

Mammalian expression vectors may also comprise an origin of replication, any necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences and 5' flanking nontranscribed sequences. In some

aspects, DNA sequences derived from the SV40 splice and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

Vectors for expressing the polypeptide or fragment thereof in eukaryotic cells may also contain enhancers to increase expression levels. Enhancers are cis-acting elements of DNA, usually from about 10 to about 300 bp in length that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin and the adenovirus enhancers.

In addition, the expression vectors typically contain one or more selectable marker genes to permit selection of host cells containing the vector. Such selectable markers include genes encoding dihydrofolate reductase or genes conferring neomycin resistance for eukaryotic cell culture, genes conferring tetracycline or ampicillin resistance in *E. coli* and the *S. cerevisiae* *TRP1* gene.

In some aspects, the nucleic acid encoding one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least about 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof is assembled in appropriate phase with a leader sequence capable of directing secretion of the translated polypeptide or fragment thereof. Optionally, the nucleic acid can encode a fusion polypeptide in which one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof is fused to heterologous peptides or polypeptides, such as N-terminal identification peptides which impart desired characteristics, such as increased stability or simplified purification.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is ligated to the desired position in the vector following digestion of the insert and the vector with appropriate restriction endonucleases. Alternatively, blunt ends in both the insert and the vector may be ligated. A variety of cloning techniques are disclosed in Ausubel *et al.* Current Protocols in Molecular Biology, John Wiley 503 Sons, Inc. 1997 and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor Laboratory Press (1989). Such procedures and others are deemed to be within the scope of those skilled in the art.

The vector may be, for example, in the form of a plasmid, a viral particle, or a phage. Other vectors include chromosomal, nonchromosomal and synthetic DNA sequences, derivatives of SV40; bacterial plasmids, phage DNA, baculovirus, yeast plasmids, vectors

derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus and pseudorabies. A variety of cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, N.Y., (1989).

5 *Host cells and transformed cells*

The invention also provides a transformed cell comprising a nucleic acid sequence of the invention, e.g., a sequence encoding a xylanase of the invention, or a vector of the invention. The host cell may be any of the host cells familiar to those skilled in the art, including prokaryotic cells, eukaryotic cells, such as bacterial cells, fungal cells, yeast cells, 10 mammalian cells, insect cells, or plant cells. Exemplary bacterial cells include *E. coli*, *Streptomyces*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*. Exemplary insect cells include *Drosophila S2* and *Spodoptera Sf9*. Exemplary animal cells include CHO, COS or Bowes melanoma or any mouse or human cell line. The selection of an appropriate host is within the 15 abilities of those skilled in the art. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, e.g., Weising (1988) Ann. Rev. Genet. 22:421-477; U.S. Patent No. 5,750,870.

The vector can be introduced into the host cells using any of a variety of techniques, including transformation, transfection, transduction, viral infection, gene guns, or 20 Ti-mediated gene transfer. Particular methods include calcium phosphate transfection, DEAE-Dextran mediated transfection, lipofection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

In one aspect, the nucleic acids or vectors of the invention are introduced into the cells for screening, thus, the nucleic acids enter the cells in a manner suitable for 25 subsequent expression of the nucleic acid. The method of introduction is largely dictated by the targeted cell type. Exemplary methods include CaPO₄ precipitation, liposome fusion, lipofection (e.g., LIPOFECTIN™), electroporation, viral infection, etc. The candidate nucleic acids may stably integrate into the genome of the host cell (for example, with retroviral introduction) or may exist either transiently or stably in the cytoplasm (i.e. through 30 the use of traditional plasmids, utilizing standard regulatory sequences, selection markers, etc.). As many pharmaceutically important screens require human or model mammalian cell targets, retroviral vectors capable of transfecting such targets are can be used.

Where appropriate, the engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the invention. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter may be induced by appropriate means (e.g., temperature shift or chemical induction) and the cells may be cultured for an additional period to allow them to produce the desired polypeptide or fragment thereof.

Cells can be harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract is retained for further purification. Microbial cells employed for expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents. Such methods are well known to those skilled in the art. The expressed polypeptide or fragment thereof can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the polypeptide. If desired, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Depending upon the host employed in a recombinant production procedure, the polypeptides produced by host cells containing the vector may be glycosylated or may be non-glycosylated. Polypeptides of the invention may or may not also include an initial methionine amino acid residue.

Cell-free translation systems can also be employed to produce a polypeptide of the invention. Cell-free translation systems can use mRNAs transcribed from a DNA construct comprising a promoter operably linked to a nucleic acid encoding the polypeptide or fragment thereof. In some aspects, the DNA construct may be linearized prior to conducting an in vitro transcription reaction. The transcribed mRNA is then incubated with an appropriate cell-free translation extract, such as a rabbit reticulocyte extract, to produce the desired polypeptide or fragment thereof.

The expression vectors can contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate

reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

Host cells containing the polynucleotides of interest, e.g., nucleic acids of the invention, can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression and will be apparent to the ordinarily skilled artisan. The clones which are identified as having the specified enzyme activity may then be sequenced to identify the polynucleotide sequence encoding an enzyme having the enhanced activity.

The invention provides a method for overexpressing a recombinant xylanase in a cell comprising expressing a vector comprising a nucleic acid of the invention, e.g., a nucleic acid comprising a nucleic acid sequence with at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to a sequence of Group A nucleic acid sequences over a region of at least about 100 residues, wherein the sequence identities are determined by analysis with a sequence comparison algorithm or by visual inspection, or, a nucleic acid that hybridizes under stringent conditions to a nucleic acid sequence as set forth in Group A nucleic acid sequences, or a subsequence thereof. The overexpression can be effected by any means, e.g., use of a high activity promoter, a dicistronic vector or by gene amplification of the vector.

The nucleic acids of the invention can be expressed, or overexpressed, in any in vitro or in vivo expression system. Any cell culture systems can be employed to express, or over-express, recombinant protein, including bacterial, insect, yeast, fungal or mammalian cultures. Over-expression can be effected by appropriate choice of promoters, enhancers, vectors (e.g., use of replicon vectors, dicistronic vectors (see, e.g., Gurtu (1996) *Biochem. Biophys. Res. Commun.* 229:295-8), media, culture systems and the like. In one aspect, gene amplification using selection markers, e.g., glutamine synthetase (see, e.g., Sanders (1987) *Dev. Biol. Stand.* 66:55-63), in cell systems are used to overexpress the polypeptides of the invention.

Additional details regarding this approach are in the public literature and/or are known to the skilled artisan. In a particular non-limiting exemplification, such publicly available literature includes EP 0659215 (W0 9403612 A1) (Nevalainen *et al.*); Lapidot, A., Mechaly, A., Shoham, Y., "Overexpression and single-step purification of a thermostable

xylanase from *Bacillus stearothermophilus* T-6," J. Biotechnol. Nov 51:259-64 (1996); Lüthi, E., Jasmat, N.B., Bergquist, P.L., "Xylanase from the extremely thermophilic bacterium *Caldocellum saccharolyticum*: overexpression of the gene in *Escherichia coli* and characterization of the gene product," Appl. Environ. Microbiol. Sep 56:2677-83 (1990); and
5 Sung, W.L., Luk, C.K., Zahab, D.M., Wakarchuk, W., "Overexpression of the *Bacillus subtilis* and *circulans* xylanases in *Escherichia coli*," Protein Expr. Purif. Jun 4:200-6 (1993), although these references do not teach the inventive enzymes of the instant application.

The host cell may be any of the host cells familiar to those skilled in the art, including prokaryotic cells, eukaryotic cells, mammalian cells, insect cells, or plant cells. As
10 representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*, *Streptomyces*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces* and *Staphylococcus*, fungal cells, such as yeast, insect cells such as *Drosophila S2* and *Spodoptera Sf9*, animal cells such as CHO, COS or Bowes melanoma and adenoviruses. The selection of an appropriate host is within the
15 abilities of those skilled in the art.

The vector may be introduced into the host cells using any of a variety of techniques, including transformation, transfection, transduction, viral infection, gene guns, or Ti-mediated gene transfer. Particular methods include calcium phosphate transfection, DEAE-Dextran mediated transfection, lipofection, or electroporation (Davis, L., Dibner, M.,
20 Battey, I., Basic Methods in Molecular Biology, (1986)).

Where appropriate, the engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the invention. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter may be induced
25 by appropriate means (*e.g.*, temperature shift or chemical induction) and the cells may be cultured for an additional period to allow them to produce the desired polypeptide or fragment thereof.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means and the resulting crude extract is retained for further purification. Microbial
30 cells employed for expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents. Such methods are well known to those skilled in the art. The expressed polypeptide or fragment thereof can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation

exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the polypeptide. If desired, high performance liquid chromatography (HPLC) can be employed for final purification steps.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts (described by Gluzman, *Cell*, 23:175, 1981) and other cell lines capable of expressing proteins from a compatible vector, such as the C127, 3T3, CHO, HeLa and BHK cell lines.

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Depending upon the host employed in a recombinant production procedure, the polypeptides produced by host cells containing the vector may be glycosylated or may be non-glycosylated. Polypeptides of the invention may or may not also include an initial methionine amino acid residue.

Alternatively, the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof can be synthetically produced by conventional peptide synthesizers. In other aspects, fragments or portions of the polypeptides may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides.

Cell-free translation systems can also be employed to produce one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof using mRNAs transcribed from a DNA construct comprising a promoter operably linked to a nucleic acid encoding the polypeptide or fragment thereof. In some aspects, the DNA construct may be linearized prior to conducting an *in vitro* transcription reaction. The transcribed mRNA is then incubated with an appropriate cell-free translation extract, such as a rabbit reticulocyte extract, to produce the desired polypeptide or fragment thereof.

Amplification of Nucleic Acids

In practicing the invention, nucleic acids of the invention and nucleic acids encoding the xylanases of the invention, or modified nucleic acids of the invention, can be reproduced by amplification. Amplification can also be used to clone or modify the nucleic acids of the invention. Thus, the invention provides amplification primer sequence pairs for amplifying nucleic acids of the invention. One of skill in the art can design amplification primer sequence pairs for any part of or the full length of these sequences.

In one aspect, the invention provides a nucleic acid amplified by a primer pair of the invention, e.g., a primer pair as set forth by about the first (the 5') 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 residues of a nucleic acid of the invention, and about the first (the 5') 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 residues of the complementary strand.

The invention provides an amplification primer sequence pair for amplifying a nucleic acid encoding a polypeptide having a xylanase activity, wherein the primer pair is capable of amplifying a nucleic acid comprising a sequence of the invention, or fragments or subsequences thereof. One or each member of the amplification primer sequence pair can comprise an oligonucleotide comprising at least about 10 to 50 consecutive bases of the sequence, or about 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 consecutive bases of the sequence. The invention provides amplification primer pairs, wherein the primer pair comprises a first member having a sequence as set forth by about the first (the 5') 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 residues of a nucleic acid of the invention, and a second member having a sequence as set forth by about the first (the 5') 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 residues of the complementary strand of the first member. The invention provides xylanases generated by amplification, e.g., polymerase chain reaction (PCR), using an amplification primer pair of the invention. The invention provides methods of making a xylanase by amplification, e.g., polymerase chain reaction (PCR), using an amplification primer pair of the invention. In one aspect, the amplification primer pair amplifies a nucleic acid from a library, e.g., a gene library, such as an environmental library.

Amplification reactions can also be used to quantify the amount of nucleic acid in a sample (such as the amount of message in a cell sample), label the nucleic acid (e.g., to apply it to an array or a blot), detect the nucleic acid, or quantify the amount of a specific nucleic acid in a sample. In one aspect of the invention, message isolated from a cell or a cDNA library are amplified.

The skilled artisan can select and design suitable oligonucleotide amplification primers. Amplification methods are also well known in the art, and include, e.g., polymerase chain reaction, PCR (see, e.g., PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, ed. Innis, Academic Press, N.Y. (1990) and PCR STRATEGIES (1995), ed. Innis, Academic Press, Inc., N.Y., ligase chain reaction (LCR) (see, e.g., Wu (1989) Genomics 4:560; Landegren (1988) Science 241:1077; Barringer (1990) Gene 89:117); transcription amplification (see, e.g., Kwoh (1989) Proc. Natl. Acad. Sci. USA 86:1173); and, self-sustained sequence replication (see, e.g., Guatelli (1990) Proc. Natl. Acad. Sci. USA 87:1874); Q Beta replicase amplification (see, e.g., Smith (1997) J. Clin. Microbiol. 35:1477-1491), automated Q-beta replicase amplification assay (see, e.g., Burg (1996) Mol. Cell. Probes 10:257-271) and other RNA polymerase mediated techniques (e.g., NASBA, Cangene, Mississauga, Ontario); see also Berger (1987) Methods Enzymol. 152:307-316; Sambrook; Ausubel; U.S. Patent Nos. 4,683,195 and 4,683,202; Sooknanan (1995) Biotechnology 13:563-564.

15 Determining the degree of sequence identity

The invention provides nucleic acids comprising sequences having at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary nucleic acid of the invention (e.g., SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID

NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID
NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID
NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID
NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID
5 NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID
NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID
NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID
NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID
NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID
10 NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID
NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID
NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID
NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID
NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID
15 NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID
NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID
NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID
NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID
NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID
20 NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID
NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID
NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID
NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID
NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID
25 NO:377 or SEQ ID NO:379) over a region of at least about 50, 75, 100, 150, 200, 250, 300,
350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150,
1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 or more, residues. The invention provides
polypeptides comprising sequences having at least about 50%, 51%, 52%, 53%, 54%, 55%,
56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%,
30 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%,
88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete
(100%) sequence identity to an exemplary polypeptide of the invention. The extent of
sequence identity (homology) may be determined using any computer program and

associated parameters, including those described herein, such as BLAST 2.2.2. or FASTA version 3.0t78, with the default parameters.

The nucleic acid sequences are also referred to as "Group A" nucleic acid sequences, which include sequences substantially identical thereto, as well as sequences homologous to Group A nucleic acid sequences and fragments thereof and sequences complementary to all of the preceding sequences. Nucleic acid sequences of the invention can comprise at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive nucleotides of an exemplary sequence of the invention (e.g., Group A nucleic acid sequences) and sequences substantially identical thereto. Homologous sequences and fragments of Group A nucleic acid sequences and sequences substantially identical thereto, refer to a sequence having at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, or 50% homology to these sequences. Homology may be determined using any of the computer programs and parameters described herein, including FASTA version 3.0t78 with the default parameters. Homologous sequences also include RNA sequences in which uridines replace the thymines in the nucleic acid sequences as set forth in the Group A nucleic acid sequences. The homologous sequences may be obtained using any of the procedures described herein or may result from the correction of a sequencing error. It will be appreciated that the nucleic acid sequences as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, can be represented in the traditional single character format (See the inside back cover of Stryer, Lubert. Biochemistry, 3rd Ed., W. H. Freeman & Co., New York.) or in any other format which records the identity of the nucleotides in a sequence.

Various sequence comparison programs identified elsewhere in this patent specification are particularly contemplated for use in this aspect of the invention. Protein and/or nucleic acid sequence homologies may be evaluated using any of the variety of sequence comparison algorithms and programs known in the art. Such algorithms and programs include, but are by no means limited to, TBLASTN, BLASTP, FASTA, TFASTA and CLUSTALW (Pearson and Lipman, Proc. Natl. Acad. Sci. USA 85(8):2444-2448, 1988; Altschul *et al.*, J. Mol. Biol. 215(3):403-410, 1990; Thompson *et al.*, Nucleic Acids Res. 22(2):4673-4680, 1994; Higgins *et al.*, Methods Enzymol. 266:383-402, 1996; Altschul *et al.*, J. Mol. Biol. 215(3):403-410, 1990; Altschul *et al.*, Nature Genetics 3:266-272, 1993).

Homology or identity is often measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Such software matches

similar sequences by assigning degrees of homology to various deletions, substitutions and other modifications. The terms "homology" and "identity" in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same when compared and aligned for maximum correspondence over a comparison window or designated region as measured using any number of sequence comparison algorithms or by manual alignment and visual inspection.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequence for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482, 1981, by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443, 1970, by the search for similarity method of person & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444, 1988, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection. Other algorithms for determining homology or identity include, for example, in addition to a BLAST program (Basic Local Alignment Search Tool at the National Center for Biological Information), ALIGN, AMAS (Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment), ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN (Biological Sequence Comparative Analysis Node), BLIMPS (BLOCKS IMPROVED Searcher), FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, Smith-Waterman algorithm, DARWIN, Las Vegas algorithm, FNAT (Forced Nucleotide Alignment Tool), Framealign, Framesearch,

DYNAMIC, FILTER, FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL, GIBBS, GenQuest, ISSC (Sensitive Sequence Comparison), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construction & Analysis Workbench), MAP (Multiple Alignment Program),

5 MBLKP, MBLKN, PIMA (Pattern-Induced Multi-sequence Alignment), SAGA (Sequence Alignment by Genetic Algorithm) and WHAT-IF. Such alignment programs can also be used to screen genome databases to identify polynucleotide sequences having substantially identical sequences. A number of genome databases are available, for example, a substantial portion of the human genome is available as part of the Human Genome Sequencing Project (J.

10 Roach, http://weber.u.Washington.edu/~roach/human_genome_progress_2.html) (Gibbs, 1995). At least twenty-one other genomes have already been sequenced, including, for example, *M. genitalium* (Fraser *et al.*, 1995), *M. jannaschii* (Bult *et al.*, 1996), *H. influenzae* (Fleischmann *et al.*, 1995), *E. coli* (Blattner *et al.*, 1997) and yeast (*S. cerevisiae*) (Mewes *et al.*, 1997) and *D. melanogaster* (Adams *et al.*, 2000). Significant progress has also been made in sequencing the

15 genomes of model organism, such as mouse, *C. elegans* and *Arabidopsis sp.* Several databases containing genomic information annotated with some functional information are maintained by different organization and are accessible via the internet

One example of a useful algorithm is BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, Nuc. Acids Res. 25:3389-3402, 1977 and Altschul *et al.*, J. Mol. Biol. 215:403-410, 1990, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is

25 referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always

30 >0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X

determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3 and expectations (E) of 10 and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989) alignments (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, Proc. Natl. Acad. Sci. USA 90:5873, 1993). One measure of similarity provided by BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a references sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01 and most preferably less than about 0.001.

In one aspect, protein and nucleic acid sequence homologies are evaluated using the Basic Local Alignment Search Tool ("BLAST") In particular, five specific BLAST programs are used to perform the following task:

(1) BLASTP and BLAST3 compare an amino acid query sequence against a protein sequence database;

(2) BLASTN compares a nucleotide query sequence against a nucleotide sequence database;

(3) BLASTX compares the six-frame conceptual translation products of a query nucleotide sequence (both strands) against a protein sequence database;

(4) TBLASTN compares a query protein sequence against a nucleotide sequence database translated in all six reading frames (both strands); and

(5) TBLASTX compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database.

The BLAST programs identify homologous sequences by identifying similar segments, which are referred to herein as "high-scoring segment pairs," between a query amino or nucleic acid sequence and a test sequence which is preferably obtained from a protein or nucleic acid sequence database. High-scoring segment pairs are preferably identified (*i.e.*, aligned) by means of a scoring matrix, many of which are known in the art. Preferably, the scoring matrix used is the BLOSUM62 matrix (Gonnet *et al.*, Science 256:1443-1445, 1992; Henikoff and Henikoff, Proteins 17:49-61, 1993). Less preferably, the

PAM or PAM250 matrices may also be used (see, e.g., Schwartz and Dayhoff, eds., 1978, *Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure*, Washington: National Biomedical Research Foundation). BLAST programs are accessible through the U.S. National Library of Medicine.

5 The parameters used with the above algorithms may be adapted depending on the sequence length and degree of homology studied. In some aspects, the parameters may be the default parameters used by the algorithms in the absence of instructions from the user.

Computer systems and computer program products

10 To determine and identify sequence identities, structural homologies, motifs and the like in silico, a nucleic acid or polypeptide sequence of the invention can be stored, recorded, and manipulated on any medium which can be read and accessed by a computer.

 Accordingly, the invention provides computers, computer systems, computer readable mediums, computer programs products and the like recorded or stored thereon the nucleic acid and polypeptide sequences of the invention. As used herein, the words "recorded"
15 and "stored" refer to a process for storing information on a computer medium. A skilled artisan can readily adopt any known methods for recording information on a computer readable medium to generate manufactures comprising one or more of the nucleic acid and/or polypeptide sequences of the invention.

 The polypeptides of the invention include the polypeptide sequences of Group
20 B amino acid sequences, the exemplary sequences of the invention, and sequences substantially identical thereto, and fragments of any of the preceding sequences. Substantially identical, or homologous, polypeptide sequences refer to a polypeptide sequence having at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%,
25 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity to an exemplary sequence of the invention, e.g., a polypeptide sequences of the Group B amino acid sequences.

 Homology may be determined using any of the computer programs and
30 parameters described herein, including FASTA version 3.0t78 with the default parameters or with any modified parameters. The homologous sequences may be obtained using any of the procedures described herein or may result from the correction of a sequencing error. The polypeptide fragments comprise at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150,

200, 250, 300, 350, 400, 450, 500 or more consecutive amino acids of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto. It will be appreciated that the polypeptide codes as set forth in Group B amino acid sequences and sequences substantially identical thereto, can be represented in the traditional single character format or three letter format (See the inside back cover of Stryer, Lubert. Biochemistry, 3rd Ed., W. H. Freeman & Co., New York.) or in any other format which relates the identity of the polypeptides in a sequence.

A nucleic acid or polypeptide sequence of the invention can be stored, recorded and manipulated on any medium which can be read and accessed by a computer. As used herein, the words "recorded" and "stored" refer to a process for storing information on a computer medium. A skilled artisan can readily adopt any of the presently known methods for recording information on a computer readable medium to generate manufactures comprising one or more of the nucleic acid sequences as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, one or more of the polypeptide sequences as set forth in Group B amino acid sequences and sequences substantially identical thereto. Another aspect of the invention is a computer readable medium having recorded thereon at least 2, 5, 10, 15, or 20 or more nucleic acid sequences as set forth in Group A nucleic acid sequences and sequences substantially identical thereto.

Another aspect of the invention is a computer readable medium having recorded thereon one or more of the nucleic acid sequences as set forth in Group A nucleic acid sequences and sequences substantially identical thereto. Another aspect of the invention is a computer readable medium having recorded thereon one or more of the polypeptide sequences as set forth in Group B amino acid sequences and sequences substantially identical thereto. Another aspect of the invention is a computer readable medium having recorded thereon at least 2, 5, 10, 15, or 20 or more of the sequences as set forth above.

Computer readable media include magnetically readable media, optically readable media, electronically readable media and magnetic/optical media. For example, the computer readable media may be a hard disk, a floppy disk, a magnetic tape, CD-ROM, Digital Versatile Disk (DVD), Random Access Memory (RAM), or Read Only Memory (ROM) as well as other types of other media known to those skilled in the art.

Aspects of the invention include systems (*e.g.*, internet based systems), particularly computer systems which store and manipulate the sequence information described herein. One example of a computer system 100 is illustrated in block diagram form in Figure 1. As used herein, "a computer system" refers to the hardware components, software components

and data storage components used to analyze a nucleotide sequence of a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in the Group B amino acid sequences. The computer system 100 typically includes a processor for processing, accessing and manipulating the sequence data. The processor 105 can be any well-known type of central processing unit, such as, for example, the Pentium III from Intel Corporation, or similar processor from Sun, Motorola, Compaq, AMD or International Business Machines.

Typically the computer system 100 is a general purpose system that comprises the processor 105 and one or more internal data storage components 110 for storing data and one or more data retrieving devices for retrieving the data stored on the data storage components. A skilled artisan can readily appreciate that any one of the currently available computer systems are suitable.

In one particular aspect, the computer system 100 includes a processor 105 connected to a bus which is connected to a main memory 115 (preferably implemented as RAM) and one or more internal data storage devices 110, such as a hard drive and/or other computer readable media having data recorded thereon. In some aspects, the computer system 100 further includes one or more data retrieving device 118 for reading the data stored on the internal data storage devices 110.

The data retrieving device 118 may represent, for example, a floppy disk drive, a compact disk drive, a magnetic tape drive, or a modem capable of connection to a remote data storage system (e.g., via the internet) etc. In some aspects, the internal data storage device 110 is a removable computer readable medium such as a floppy disk, a compact disk, a magnetic tape, etc. containing control logic and/or data recorded thereon. The computer system 100 may advantageously include or be programmed by appropriate software for reading the control logic and/or the data from the data storage component once inserted in the data retrieving device.

The computer system 100 includes a display 120 which is used to display output to a computer user. It should also be noted that the computer system 100 can be linked to other computer systems 125a-c in a network or wide area network to provide centralized access to the computer system 100.

Software for accessing and processing the nucleotide sequences of a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, (such as search tools, compare tools and modeling tools etc.) may reside in main memory 115 during execution.

In some aspects, the computer system 100 may further comprise a sequence comparison algorithm for comparing a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, stored on a computer readable medium to a reference nucleotide or polypeptide sequence(s) stored on a computer readable medium. A "sequence comparison algorithm" refers to one or more programs which are implemented (locally or remotely) on the computer system 100 to compare a nucleotide sequence with other nucleotide sequences and/or compounds stored within a data storage means. For example, the sequence comparison algorithm may compare the nucleotide sequences of a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, stored on a computer readable medium to reference sequences stored on a computer readable medium to identify homologies or structural motifs.

Figure 2 is a flow diagram illustrating one aspect of a process 200 for comparing a new nucleotide or protein sequence with a database of sequences in order to determine the homology levels between the new sequence and the sequences in the database. The database of sequences can be a private database stored within the computer system 100, or a public database such as GENBANK that is available through the Internet.

The process 200 begins at a start state 201 and then moves to a state 202 wherein the new sequence to be compared is stored to a memory in a computer system 100. As discussed above, the memory could be any type of memory, including RAM or an internal storage device.

The process 200 then moves to a state 204 wherein a database of sequences is opened for analysis and comparison. The process 200 then moves to a state 206 wherein the first sequence stored in the database is read into a memory on the computer. A comparison is then performed at a state 210 to determine if the first sequence is the same as the second sequence. It is important to note that this step is not limited to performing an exact comparison between the new sequence and the first sequence in the database. Well-known methods are known to those of skill in the art for comparing two nucleotide or protein sequences, even if they are not identical. For example, gaps can be introduced into one sequence in order to raise the homology level between the two tested sequences. The parameters that control whether gaps or other features are introduced into a sequence during comparison are normally entered by the user of the computer system.

Once a comparison of the two sequences has been performed at the state 210, a determination is made at a decision state 210 whether the two sequences are the same. Of course, the term "same" is not limited to sequences that are absolutely identical. Sequences that are within the homology parameters entered by the user will be marked as "same" in the process 200.

If a determination is made that the two sequences are the same, the process 200 moves to a state 214 wherein the name of the sequence from the database is displayed to the user. This state notifies the user that the sequence with the displayed name fulfills the homology constraints that were entered. Once the name of the stored sequence is displayed to the user, the process 200 moves to a decision state 218 wherein a determination is made whether more sequences exist in the database. If no more sequences exist in the database, then the process 200 terminates at an end state 220. However, if more sequences do exist in the database, then the process 200 moves to a state 224 wherein a pointer is moved to the next sequence in the database so that it can be compared to the new sequence. In this manner, the new sequence is aligned and compared with every sequence in the database.

It should be noted that if a determination had been made at the decision state 212 that the sequences were not homologous, then the process 200 would move immediately to the decision state 218 in order to determine if any other sequences were available in the database for comparison.

Accordingly, one aspect of the invention is a computer system comprising a processor, a data storage device having stored thereon a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, a data storage device having retrievably stored thereon reference nucleotide sequences or polypeptide sequences to be compared to a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto and a sequence comparer for conducting the comparison. The sequence comparer may indicate a homology level between the sequences compared or identify structural motifs in the above described nucleic acid code of Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, or it may identify structural motifs in sequences which are compared to these nucleic acid codes and polypeptide codes. In some aspects, the data storage device may have stored thereon the

sequences of at least 2, 5, 10, 15, 20, 25, 30 or 40 or more of the nucleic acid sequences as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or the polypeptide sequences as set forth in Group B amino acid sequences and sequences substantially identical thereto.

5 Another aspect of the invention is a method for determining the level of homology between a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto and a reference nucleotide sequence. The method including reading the nucleic acid code or the polypeptide code and the
10 reference nucleotide or polypeptide sequence through the use of a computer program which determines homology levels and determining homology between the nucleic acid code or polypeptide code and the reference nucleotide or polypeptide sequence with the computer program. The computer program may be any of a number of computer programs for determining homology levels, including those specifically enumerated herein, (e.g., BLAST2N
15 with the default parameters or with any modified parameters). The method may be implemented using the computer systems described above. The method may also be performed by reading at least 2, 5, 10, 15, 20, 25, 30 or 40 or more of the above described nucleic acid sequences as set forth in the Group A nucleic acid sequences, or the polypeptide sequences as set forth in the Group B amino acid sequences through use of the computer program and determining
20 homology between the nucleic acid codes or polypeptide codes and reference nucleotide sequences or polypeptide sequences.

Figure 3 is a flow diagram illustrating one aspect of a process 250 in a computer for determining whether two sequences are homologous. The process 250 begins at a start state 252 and then moves to a state 254 wherein a first sequence to be compared is
25 stored to a memory. The second sequence to be compared is then stored to a memory at a state 256. The process 250 then moves to a state 260 wherein the first character in the first sequence is read and then to a state 262 wherein the first character of the second sequence is read. It should be understood that if the sequence is a nucleotide sequence, then the character would normally be either A, T, C, G or U. If the sequence is a protein sequence, then it is
30 preferably in the single letter amino acid code so that the first and sequence sequences can be easily compared.

A determination is then made at a decision state 264 whether the two characters are the same. If they are the same, then the process 250 moves to a state 268 wherein the next characters in the first and second sequences are read. A determination is

then made whether the next characters are the same. If they are, then the process 250 continues this loop until two characters are not the same. If a determination is made that the next two characters are not the same, the process 250 moves to a decision state 274 to determine whether there are any more characters either sequence to read.

5 If there are not any more characters to read, then the process 250 moves to a state 276 wherein the level of homology between the first and second sequences is displayed to the user. The level of homology is determined by calculating the proportion of characters between the sequences that were the same out of the total number of sequences in the first sequence. Thus, if every character in a first 100 nucleotide sequence aligned with a every
10 character in a second sequence, the homology level would be 100%.

Alternatively, the computer program may be a computer program which compares the nucleotide sequences of a nucleic acid sequence as set forth in the invention, to one or more reference nucleotide sequences in order to determine whether the nucleic acid code of Group A nucleic acid sequences and sequences substantially identical thereto, differs from
15 a reference nucleic acid sequence at one or more positions. Optionally such a program records the length and identity of inserted, deleted or substituted nucleotides with respect to the sequence of either the reference polynucleotide or a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto. In one aspect, the computer program may be a program which determines whether a nucleic acid sequence as set forth in
20 Group A nucleic acid sequences and sequences substantially identical thereto, contains a single nucleotide polymorphism (SNP) with respect to a reference nucleotide sequence.

Accordingly, another aspect of the invention is a method for determining whether a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, differs at one or more nucleotides from a reference
25 nucleotide sequence comprising the steps of reading the nucleic acid code and the reference nucleotide sequence through use of a computer program which identifies differences between nucleic acid sequences and identifying differences between the nucleic acid code and the reference nucleotide sequence with the computer program. In some aspects, the computer program is a program which identifies single nucleotide polymorphisms. The method may be
30 implemented by the computer systems described above and the method illustrated in Figure 3. The method may also be performed by reading at least 2, 5, 10, 15, 20, 25, 30, or 40 or more of the nucleic acid sequences as set forth in Group A nucleic acid sequences and sequences substantially identical thereto and the reference nucleotide sequences through the use of the

computer program and identifying differences between the nucleic acid codes and the reference nucleotide sequences with the computer program.

In other aspects the computer based system may further comprise an identifier for identifying features within a nucleic acid sequence as set forth in the Group A nucleic acid sequences or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto.

An "identifier" refers to one or more programs which identifies certain features within a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto. In one aspect, the identifier may comprise a program which identifies an open reading frame in a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto.

Figure 4 is a flow diagram illustrating one aspect of an identifier process 300 for detecting the presence of a feature in a sequence. The process 300 begins at a start state 302 and then moves to a state 304 wherein a first sequence that is to be checked for features is stored to a memory 115 in the computer system 100. The process 300 then moves to a state 306 wherein a database of sequence features is opened. Such a database would include a list of each feature's attributes along with the name of the feature. For example, a feature name could be "Initiation Codon" and the attribute would be "ATG". Another example would be the feature name "TAATAA Box" and the feature attribute would be "TAATAA". An example of such a database is produced by the University of Wisconsin Genetics Computer Group. Alternatively, the features may be structural polypeptide motifs such as alpha helices, beta sheets, or functional polypeptide motifs such as enzymatic active sites, helix-turn-helix motifs or other motifs known to those skilled in the art.

Once the database of features is opened at the state 306, the process 300 moves to a state 308 wherein the first feature is read from the database. A comparison of the attribute of the first feature with the first sequence is then made at a state 310. A determination is then made at a decision state 316 whether the attribute of the feature was found in the first sequence. If the attribute was found, then the process 300 moves to a state 318 wherein the name of the found feature is displayed to the user.

The process 300 then moves to a decision state 320 wherein a determination is made whether more features exist in the database. If no more features do exist, then the process 300 terminates at an end state 324. However, if more features do exist in the

database, then the process 300 reads the next sequence feature at a state 326 and loops back to the state 310 wherein the attribute of the next feature is compared against the first sequence. It should be noted, that if the feature attribute is not found in the first sequence at the decision state 316, the process 300 moves directly to the decision state 320 in order to
5 determine if any more features exist in the database.

Accordingly, another aspect of the invention is a method of identifying a feature within a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, comprising reading the
10 nucleic acid code(s) or polypeptide code(s) through the use of a computer program which identifies features therein and identifying features within the nucleic acid code(s) with the computer program. In one aspect, computer program comprises a computer program which identifies open reading frames. The method may be performed by reading a single sequence or at least 2, 5, 10, 15, 20, 25, 30, or 40 of the nucleic acid sequences as set forth in Group A
15 nucleic acid sequences and sequences substantially identical thereto, or the polypeptide sequences as set forth in Group B amino acid sequences and sequences substantially identical thereto, through the use of the computer program and identifying features within the nucleic acid codes or polypeptide codes with the computer program.

A nucleic acid sequence as set forth in Group A nucleic acid sequences and
20 sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto, may be stored and manipulated in a variety of data processor programs in a variety of formats. For example, a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid
25 sequences and sequences substantially identical thereto, may be stored as text in a word processing file, such as Microsoft WORD™ or WORDPERFECT™ or as an ASCII file in a variety of database programs familiar to those of skill in the art, such as DB2™, SYBASE™, or ORACLE™. In addition, many computer programs and databases may be used as sequence comparison algorithms, identifiers, or sources of reference nucleotide sequences or polypeptide
30 sequences to be compared to a nucleic acid sequence as set forth in Group A nucleic acid sequences and sequences substantially identical thereto, or a polypeptide sequence as set forth in Group B amino acid sequences and sequences substantially identical thereto. The following list is intended not to limit the invention but to provide guidance to programs and databases which are useful with the nucleic acid sequences as set forth in Group A nucleic acid

sequences and sequences substantially identical thereto, or the polypeptide sequences as set forth in Group B amino acid sequences and sequences substantially identical thereto.

The programs and databases which may be used include, but are not limited to: MacPattern (EMBL), DiscoveryBase (Molecular Applications Group), GeneMine (Molecular Applications Group), Look (Molecular Applications Group), MacLook (Molecular Applications Group), BLAST and BLAST2 (NCBI), BLASTN and BLASTX (Altschul et al, J. Mol. Biol. 215: 403, 1990), FASTA (Pearson and Lipman, Proc. Natl. Acad. Sci. USA, 85: 2444, 1988), FASTDB (Brutlag *et al.* Comp. App. Biosci. 6:237-245, 1990), Catalyst (Molecular Simulations Inc.), Catalyst/SHAPE (Molecular Simulations Inc.), Cerius².DBAccess (Molecular Simulations Inc.), HypoGen (Molecular Simulations Inc.), Insight II, (Molecular Simulations Inc.), Discover (Molecular Simulations Inc.), CHARMM (Molecular Simulations Inc.), Felix (Molecular Simulations Inc.), DelPhi, (Molecular Simulations Inc.), QuanteMM, (Molecular Simulations Inc.), Homology (Molecular Simulations Inc.), Modeler (Molecular Simulations Inc.), ISIS (Molecular Simulations Inc.), Quanta/Protein Design (Molecular Simulations Inc.), WebLab (Molecular Simulations Inc.), WebLab Diversity Explorer (Molecular Simulations Inc.), Gene Explorer (Molecular Simulations Inc.), SeqFold (Molecular Simulations Inc.), the MDL Available Chemicals Directory database, the MDL Drug Data Report data base, the Comprehensive Medicinal Chemistry database, Derwents's World Drug Index database, the BioByteMasterFile database, the Genbank database and the Genseq database. Many other programs and data bases would be apparent to one of skill in the art given the present disclosure.

Motifs which may be detected using the above programs include sequences encoding leucine zippers, helix-turn-helix motifs, glycosylation sites, ubiquitination sites, alpha helices and beta sheets, signal sequences encoding signal peptides which direct the secretion of the encoded proteins, sequences implicated in transcription regulation such as homeoboxes, acidic stretches, enzymatic active sites, substrate binding sites and enzymatic cleavage sites.

Hybridization of nucleic acids

The invention provides isolated or recombinant nucleic acids that hybridize under stringent conditions to an exemplary sequence of the invention (e.g., SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45,

SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, 5 SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID 10 NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID 15 NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID 20 NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID 25 NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID 30 NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379). The stringent conditions can be highly stringent conditions, medium stringent conditions and/or

low stringent conditions, including the high and reduced stringency conditions described herein. In one aspect, it is the stringency of the wash conditions that set forth the conditions which determine whether a nucleic acid is within the scope of the invention, as discussed below.

5 In alternative aspects, nucleic acids of the invention as defined by their ability to hybridize under stringent conditions can be between about five residues and the full length of nucleic acid of the invention; e.g., they can be at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or more, residues in length. Nucleic acids shorter than full length
10 are also included. These nucleic acids can be useful as, e.g., hybridization probes, labeling probes, PCR oligonucleotide probes, iRNA (single or double stranded), antisense or sequences encoding antibody binding peptides (epitopes), motifs, active sites and the like.

In one aspect, nucleic acids of the invention are defined by their ability to hybridize under high stringency comprising conditions of about 50% formamide at about 37°C
15 to 42°C. In one aspect, nucleic acids of the invention are defined by their ability to hybridize under reduced stringency comprising conditions in about 35% to 25% formamide at about 30°C to 35°C.

Alternatively, nucleic acids of the invention are defined by their ability to hybridize under high stringency comprising conditions at 42°C in 50% formamide, 5X SSPE,
20 0.3% SDS, and a repetitive sequence blocking nucleic acid, such as cot-1 or salmon sperm DNA (e.g., 200 n/ml sheared and denatured salmon sperm DNA). In one aspect, nucleic acids of the invention are defined by their ability to hybridize under reduced stringency conditions comprising 35% formamide at a reduced temperature of 35°C.

In nucleic acid hybridization reactions, the conditions used to achieve a
25 particular level of stringency will vary, depending on the nature of the nucleic acids being hybridized. For example, the length, degree of complementarity, nucleotide sequence composition (e.g., GC v. AT content) and nucleic acid type (e.g., RNA v. DNA) of the hybridizing regions of the nucleic acids can be considered in selecting hybridization conditions. An additional consideration is whether one of the nucleic acids is immobilized,
30 for example, on a filter.

Hybridization may be carried out under conditions of low stringency, moderate stringency or high stringency. As an example of nucleic acid hybridization, a polymer membrane containing immobilized denatured nucleic acids is first prehybridized for 30

minutes at 45°C in a solution consisting of 0.9 M NaCl, 50 mM NaH₂PO₄, pH 7.0, 5.0 mM Na₂EDTA, 0.5% SDS, 10X Denhardt's and 0.5 mg/ml polyriboadenylic acid. Approximately 2 X 10⁷ cpm (specific activity 4-9 X 10⁸ cpm/ug) of ³²P end-labeled oligonucleotide probe are then added to the solution. After 12-16 hours of incubation, the membrane is washed for 30 minutes at room temperature in 1X SET (150 mM NaCl, 20 mM Tris hydrochloride, pH 7.8, 1 mM Na₂EDTA) containing 0.5% SDS, followed by a 30 minute wash in fresh 1X SET at T_m-10°C for the oligonucleotide probe. The membrane is then exposed to autoradiographic film for detection of hybridization signals.

All of the foregoing hybridizations would be considered to be under conditions of high stringency.

Following hybridization, a filter can be washed to remove any non-specifically bound detectable probe. The stringency used to wash the filters can also be varied depending on the nature of the nucleic acids being hybridized, the length of the nucleic acids being hybridized, the degree of complementarity, the nucleotide sequence composition (*e.g.*, GC v. AT content) and the nucleic acid type (*e.g.*, RNA v. DNA). Examples of progressively higher stringency condition washes are as follows: 2X SSC, 0.1% SDS at room temperature for 15 minutes (low stringency); 0.1X SSC, 0.5% SDS at room temperature for 30 minutes to 1 hour (moderate stringency); 0.1X SSC, 0.5% SDS for 15 to 30 minutes at between the hybridization temperature and 68°C (high stringency); and 0.15M NaCl for 15 minutes at 72°C (very high stringency). A final low stringency wash can be conducted in 0.1X SSC at room temperature. The examples above are merely illustrative of one set of conditions that can be used to wash filters. One of skill in the art would know that there are numerous recipes for different stringency washes. Some other examples are given below.

Nucleic acids which have hybridized to the probe are identified by autoradiography or other conventional techniques.

The above procedure may be modified to identify nucleic acids having decreasing levels of homology to the probe sequence. For example, to obtain nucleic acids of decreasing homology to the detectable probe, less stringent conditions may be used. For example, the hybridization temperature may be decreased in increments of 5°C from 68°C to 42°C in a hybridization buffer having a Na⁺ concentration of approximately 1M. Following hybridization, the filter may be washed with 2X SSC, 0.5% SDS at the temperature of hybridization. These conditions are considered to be "moderate" conditions above 50°C and "low" conditions below 50°C. A specific example of "moderate" hybridization conditions is

when the above hybridization is conducted at 55°C. A specific example of "low stringency" hybridization conditions is when the above hybridization is conducted at 45°C.

Alternatively, the hybridization may be carried out in buffers, such as 6X SSC, containing formamide at a temperature of 42°C. In this case, the concentration of formamide in the hybridization buffer may be reduced in 5% increments from 50% to 0% to identify clones having decreasing levels of homology to the probe. Following hybridization, the filter may be washed with 6X SSC, 0.5% SDS at 50°C. These conditions are considered to be "moderate" conditions above 25% formamide and "low" conditions below 25% formamide. A specific example of "moderate" hybridization conditions is when the above hybridization is conducted at 30% formamide. A specific example of "low stringency" hybridization conditions is when the above hybridization is conducted at 10% formamide.

However, the selection of a hybridization format is not critical - it is the stringency of the wash conditions that set forth the conditions which determine whether a nucleic acid is within the scope of the invention. Wash conditions used to identify nucleic acids within the scope of the invention include, e.g.: a salt concentration of about 0.02 molar at pH 7 and a temperature of at least about 50°C or about 55°C to about 60°C; or, a salt concentration of about 0.15 M NaCl at 72°C for about 15 minutes; or, a salt concentration of about 0.2X SSC at a temperature of at least about 50°C or about 55°C to about 60°C for about 15 to about 20 minutes; or, the hybridization complex is washed twice with a solution with a salt concentration of about 2X SSC containing 0.1% SDS at room temperature for 15 minutes and then washed twice by 0.1X SSC containing 0.1% SDS at 68°C for 15 minutes; or, equivalent conditions. See Sambrook, Tijssen and Ausubel for a description of SSC buffer and equivalent conditions.

These methods may be used to isolate nucleic acids of the invention. For example, the preceding methods may be used to isolate nucleic acids having a sequence with at least about 97%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, or at least 50% homology to a nucleic acid sequence selected from the group consisting of one of the sequences of Group A nucleic acid sequences and sequences substantially identical thereto, or fragments comprising at least about 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive bases thereof and the sequences complementary thereto. Homology may be measured using the alignment algorithm. For example, the homologous polynucleotides may have a coding sequence which is a naturally occurring allelic variant of one of the coding sequences

described herein. Such allelic variants may have a substitution, deletion or addition of one or more nucleotides when compared to the nucleic acids of Group A nucleic acid sequences or the sequences complementary thereto.

5 Additionally, the above procedures may be used to isolate nucleic acids which encode polypeptides having at least about 99%, 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, or at least 50% homology to a polypeptide having the sequence of one of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof as determined using a sequence alignment
10 algorithm (*e.g.*, such as the FASTA version 3.0t78 algorithm with the default parameters).

Oligonucleotides probes and methods for using them

The invention also provides nucleic acid probes that can be used, *e.g.*, for identifying nucleic acids encoding a polypeptide with a xylanase activity or fragments thereof or for identifying xylanase genes. In one aspect, the probe comprises at least 10 consecutive
15 bases of a nucleic acid of the invention. Alternatively, a probe of the invention can be at least about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 150 or about 10 to 50, about 20 to 60 about 30 to 70, consecutive bases of a sequence as set forth in a nucleic acid of the invention. The probes identify a nucleic acid by binding and/or hybridization. The probes can be used in arrays of
20 the invention, see discussion below, including, *e.g.*, capillary arrays. The probes of the invention can also be used to isolate other nucleic acids or polypeptides.

The isolated nucleic acids of Group A nucleic acid sequences and sequences substantially identical thereto, the sequences complementary thereto, or a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500
25 consecutive bases of one of the sequences of Group A nucleic acid sequences and sequences substantially identical thereto, or the sequences complementary thereto may also be used as probes to determine whether a biological sample, such as a soil sample, contains an organism having a nucleic acid sequence of the invention or an organism from which the nucleic acid was obtained. In such procedures, a biological sample potentially harboring the organism
30 from which the nucleic acid was isolated is obtained and nucleic acids are obtained from the sample. The nucleic acids are contacted with the probe under conditions which permit the probe to specifically hybridize to any complementary sequences from which are present therein.

Where necessary, conditions which permit the probe to specifically hybridize to complementary sequences may be determined by placing the probe in contact with complementary sequences from samples known to contain the complementary sequence as well as control sequences which do not contain the complementary sequence. Hybridization conditions, such as the salt concentration of the hybridization buffer, the formamide concentration of the hybridization buffer, or the hybridization temperature, may be varied to identify conditions which allow the probe to hybridize specifically to complementary nucleic acids.

If the sample contains the organism from which the nucleic acid was isolated, specific hybridization of the probe is then detected. Hybridization may be detected by labeling the probe with a detectable agent such as a radioactive isotope, a fluorescent dye or an enzyme capable of catalyzing the formation of a detectable product.

Many methods for using the labeled probes to detect the presence of complementary nucleic acids in a sample are familiar to those skilled in the art. These include Southern Blots, Northern Blots, colony hybridization procedures and dot blots. Protocols for each of these procedures are provided in Ausubel *et al.* Current Protocols in Molecular Biology, John Wiley 503 Sons, Inc. (1997) and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor Laboratory Press (1989).

Alternatively, more than one probe (at least one of which is capable of specifically hybridizing to any complementary sequences which are present in the nucleic acid sample), may be used in an amplification reaction to determine whether the sample contains an organism containing a nucleic acid sequence of the invention (*e.g.*, an organism from which the nucleic acid was isolated). Typically, the probes comprise oligonucleotides. In one aspect, the amplification reaction may comprise a PCR reaction. PCR protocols are described in Ausubel and Sambrook, *supra*. Alternatively, the amplification may comprise a ligase chain reaction, 3SR, or strand displacement reaction. (See Barany, F., "The Ligase Chain Reaction in a PCR World", *PCR Methods and Applications* 1:5-16, 1991; E. Fahy *et al.*, "Self-sustained Sequence Replication (3SR): An Isothermal Transcription-based Amplification System Alternative to PCR", *PCR Methods and Applications* 1:25-33, 1991; and Walker G.T. *et al.*, "Strand Displacement Amplification-an Isothermal *in vitro* DNA Amplification Technique", *Nucleic Acid Research* 20:1691-1696, 1992). In such procedures, the nucleic acids in the sample are contacted with the probes, the amplification reaction is performed and any resulting amplification product is detected. The amplification product may be detected by performing gel electrophoresis on the reaction products and staining the gel with an intercalator such as

ethidium bromide. Alternatively, one or more of the probes may be labeled with a radioactive isotope and the presence of a radioactive amplification product may be detected by autoradiography after gel electrophoresis.

Probes derived from sequences near the ends of the sequences of Group A nucleic acid sequences and sequences substantially identical thereto, may also be used in chromosome walking procedures to identify clones containing genomic sequences located adjacent to the sequences of Group A nucleic acid sequences and sequences substantially identical thereto. Such methods allow the isolation of genes which encode additional proteins from the host organism.

The isolated nucleic acids of Group A nucleic acid sequences and sequences substantially identical thereto, the sequences complementary thereto, or a fragment comprising at least 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, or 500 consecutive bases of one of the sequences of Group A nucleic acid sequences and sequences substantially identical thereto, or the sequences complementary thereto may be used as probes to identify and isolate related nucleic acids. In some aspects, the related nucleic acids may be cDNAs or genomic DNAs from organisms other than the one from which the nucleic acid was isolated. For example, the other organisms may be related organisms. In such procedures, a nucleic acid sample is contacted with the probe under conditions which permit the probe to specifically hybridize to related sequences. Hybridization of the probe to nucleic acids from the related organism is then detected using any of the methods described above.

By varying the stringency of the hybridization conditions used to identify nucleic acids, such as cDNAs or genomic DNAs, which hybridize to the detectable probe, nucleic acids having different levels of homology to the probe can be identified and isolated. Stringency may be varied by conducting the hybridization at varying temperatures below the melting temperatures of the probes. The melting temperature, T_m , is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly complementary probe. Very stringent conditions are selected to be equal to or about 5°C lower than the T_m for a particular probe. The melting temperature of the probe may be calculated using the following formulas:

For probes between 14 and 70 nucleotides in length the melting temperature (T_m) is calculated using the formula: $T_m = 81.5 + 16.6(\log [Na^+]) + 0.41(\text{fraction G+C}) - (600/N)$ where N is the length of the probe.

If the hybridization is carried out in a solution containing formamide, the melting temperature may be calculated using the equation: $T_m = 81.5 + 16.6(\log [Na^+]) + 0.41(\text{fraction } G+C) - (0.63\% \text{ formamide}) - (600/N)$ where N is the length of the probe.

Prehybridization may be carried out in 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 μ g denatured fragmented salmon sperm DNA or 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 μ g denatured fragmented salmon sperm DNA, 50% formamide. The formulas for SSC and Denhardt's solutions are listed in Sambrook *et al.*, *supra*.

Hybridization is conducted by adding the detectable probe to the prehybridization solutions listed above. Where the probe comprises double stranded DNA, it is denatured before addition to the hybridization solution. The filter is contacted with the hybridization solution for a sufficient period of time to allow the probe to hybridize to cDNAs or genomic DNAs containing sequences complementary thereto or homologous thereto. For probes over 200 nucleotides in length, the hybridization may be carried out at 15-25°C below the T_m . For shorter probes, such as oligonucleotide probes, the hybridization may be conducted at 5-10°C below the T_m . Typically, for hybridizations in 6X SSC, the hybridization is conducted at approximately 68°C. Usually, for hybridizations in 50% formamide containing solutions, the hybridization is conducted at approximately 42°C.

Inhibiting Expression of Xylanases

The invention provides nucleic acids complementary to (e.g., antisense sequences to) the nucleic acids of the invention, e.g., xylanase-encoding nucleic acids. Antisense sequences are capable of inhibiting the transport, splicing or transcription of xylanase-encoding genes. The inhibition can be effected through the targeting of genomic DNA or messenger RNA. The transcription or function of targeted nucleic acid can be inhibited, for example, by hybridization and/or cleavage. One particularly useful set of inhibitors provided by the present invention includes oligonucleotides which are able to either bind xylanase gene or message, in either case preventing or inhibiting the production or function of xylanase. The association can be through sequence specific hybridization. Another useful class of inhibitors includes oligonucleotides which cause inactivation or cleavage of xylanase message. The oligonucleotide can have enzyme activity which causes such cleavage, such as ribozymes. The oligonucleotide can be chemically modified or conjugated to an enzyme or composition capable of cleaving the complementary nucleic acid. A pool of many different such oligonucleotides can be screened for those with the desired activity. Thus, the invention provides various compositions for the inhibition of xylanase

expression on a nucleic acid and/or protein level, e.g., antisense, iRNA and ribozymes comprising xylanase sequences of the invention and the anti-xylanase antibodies of the invention.

Inhibition of xylanase expression can have a variety of industrial applications.

5 For example, inhibition of xylanase expression can slow or prevent spoilage. Spoilage can occur when polysaccharides, e.g., structural polysaccharides, are enzymatically degraded. This can lead to the deterioration, or rot, of fruits and vegetables. In one aspect, use of compositions of the invention that inhibit the expression and/or activity of xylanases, e.g., antibodies, antisense oligonucleotides, ribozymes and RNAi, are used to slow or prevent
10 spoilage. Thus, in one aspect, the invention provides methods and compositions comprising application onto a plant or plant product (e.g., a cereal, a grain, a fruit, seed, root, leaf, etc.) antibodies, antisense oligonucleotides, ribozymes and RNAi of the invention to slow or prevent spoilage. These compositions also can be expressed by the plant (e.g., a transgenic plant) or another organism (e.g., a bacterium or other microorganism transformed with a
15 xylanase gene of the invention).

The compositions of the invention for the inhibition of xylanase expression (e.g., antisense, iRNA, ribozymes, antibodies) can be used as pharmaceutical compositions, e.g., as anti-pathogen agents or in other therapies, e.g., as anti-microbials for, e.g., *Salmonella*.

20 *Antisense Oligonucleotides*

The invention provides antisense oligonucleotides capable of binding xylanase message which can inhibit xylan hydrolase activity (e.g., catalyzing hydrolysis of internal β -1,4-xylosidic linkages) by targeting mRNA. Strategies for designing antisense oligonucleotides are well described in the scientific and patent literature, and the skilled
25 artisan can design such xylanase oligonucleotides using the novel reagents of the invention. For example, gene walking/ RNA mapping protocols to screen for effective antisense oligonucleotides are well known in the art, see, e.g., Ho (2000) *Methods Enzymol.* 314:168-183, describing an RNA mapping assay, which is based on standard molecular techniques to provide an easy and reliable method for potent antisense sequence selection. See also Smith
30 (2000) *Eur. J. Pharm. Sci.* 11:191-198.

Naturally occurring nucleic acids are used as antisense oligonucleotides. The antisense oligonucleotides can be of any length; for example, in alternative aspects, the antisense oligonucleotides are between about 5 to 100, about 10 to 80, about 15 to 60, about

18 to 40. The optimal length can be determined by routine screening. The antisense oligonucleotides can be present at any concentration. The optimal concentration can be determined by routine screening. A wide variety of synthetic, non-naturally occurring nucleotide and nucleic acid analogues are known which can address this potential problem.

5 For example, peptide nucleic acids (PNAs) containing non-ionic backbones, such as N-(2-aminoethyl) glycine units can be used. Antisense oligonucleotides having phosphorothioate linkages can also be used, as described in WO 97/03211; WO 96/39154; Mata (1997) Toxicol Appl Pharmacol 144:189-197; Antisense Therapeutics, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Antisense oligonucleotides having synthetic DNA backbone analogues provided
10 by the invention can also include phosphoro-dithioate, methylphosphonate, phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, and morpholino carbamate nucleic acids, as described above.

Combinatorial chemistry methodology can be used to create vast numbers of oligonucleotides that can be rapidly screened for specific oligonucleotides that have
15 appropriate binding affinities and specificities toward any target, such as the sense and antisense xylanase sequences of the invention (see, e.g., Gold (1995) J. of Biol. Chem. 270:13581-13584).

Inhibitory Ribozymes

The invention provides ribozymes capable of binding xylanase message.
20 These ribozymes can inhibit xylanase activity by, e.g., targeting mRNA. Strategies for designing ribozymes and selecting the xylanase-specific antisense sequence for targeting are well described in the scientific and patent literature, and the skilled artisan can design such ribozymes using the novel reagents of the invention. Ribozymes act by binding to a target RNA through the target RNA binding portion of a ribozyme which is held in close proximity
25 to an enzymatic portion of the RNA that cleaves the target RNA. Thus, the ribozyme recognizes and binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cleave and inactivate the target RNA. Cleavage of a target RNA in such a manner will destroy its ability to direct synthesis of an encoded protein if the cleavage occurs in the coding sequence. After a ribozyme has bound and cleaved its
30 RNA target, it can be released from that RNA to bind and cleave new targets repeatedly.

In some circumstances, the enzymatic nature of a ribozyme can be advantageous over other technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its transcription, translation or

association with another molecule) as the effective concentration of ribozyme necessary to effect a therapeutic treatment can be lower than that of an antisense oligonucleotide. This potential advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, a ribozyme is typically a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding, but also on the mechanism by which the molecule inhibits the expression of the RNA to which it binds. That is, the inhibition is caused by cleavage of the RNA target and so specificity is defined as the ratio of the rate of cleavage of the targeted RNA over the rate of cleavage of non-targeted RNA. This cleavage mechanism is dependent upon factors additional to those involved in base pairing. Thus, the specificity of action of a ribozyme can be greater than that of antisense oligonucleotide binding the same RNA site.

The ribozyme of the invention, e.g., an enzymatic ribozyme RNA molecule, can be formed in a hammerhead motif, a hairpin motif, as a hepatitis delta virus motif, a group I intron motif and/or an RNaseP-like RNA in association with an RNA guide sequence. Examples of hammerhead motifs are described by, e.g., Rossi (1992) Aids Research and Human Retroviruses 8:183; hairpin motifs by Hampel (1989) Biochemistry 28:4929, and Hampel (1990) Nuc. Acids Res. 18:299; the hepatitis delta virus motif by Perrotta (1992) Biochemistry 31:16; the RNaseP motif by Guerrier-Takada (1983) Cell 35:849; and the group I intron by Cech U.S. Pat. No. 4,987,071. The recitation of these specific motifs is not intended to be limiting. Those skilled in the art will recognize that a ribozyme of the invention, e.g., an enzymatic RNA molecule of this invention, can have a specific substrate binding site complementary to one or more of the target gene RNA regions. A ribozyme of the invention can have a nucleotide sequence within or surrounding that substrate binding site which imparts an RNA cleaving activity to the molecule.

RNA interference (RNAi)

In one aspect, the invention provides an RNA inhibitory molecule, a so-called "RNAi" molecule, comprising a xylanase sequence of the invention. The RNAi molecule comprises a double-stranded RNA (dsRNA) molecule. The RNAi can inhibit expression of a xylanase gene. In one aspect, the RNAi is about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more duplex nucleotides in length. While the invention is not limited by any particular mechanism of action, the RNAi can enter a cell and cause the degradation of a single-stranded RNA (ssRNA) of similar or identical sequences, including endogenous mRNAs.

When a cell is exposed to double-stranded RNA (dsRNA), mRNA from the homologous gene is selectively degraded by a process called RNA interference (RNAi). A possible basic mechanism behind RNAi is the breaking of a double-stranded RNA (dsRNA) matching a specific gene sequence into short pieces called short interfering RNA, which trigger the degradation of mRNA that matches its sequence. In one aspect, the RNAi's of the invention are used in gene-silencing therapeutics, see, e.g., Shuey (2002) Drug Discov. Today 7:1040-1046. In one aspect, the invention provides methods to selectively degrade RNA using the RNAi's of the invention. The process may be practiced *in vitro*, *ex vivo* or *in vivo*. In one aspect, the RNAi molecules of the invention can be used to generate a loss-of-function mutation in a cell, an organ or an animal. Methods for making and using RNAi molecules for selectively degrade RNA are well known in the art, see, e.g., U.S. Patent No. 6,506,559; 6,511,824; 6,515,109; 6,489,127.

Modification of Nucleic Acids

The invention provides methods of generating variants of the nucleic acids of the invention, e.g., those encoding a xylanase. These methods can be repeated or used in various combinations to generate xylanases having an altered or different activity or an altered or different stability from that of a xylanase encoded by the template nucleic acid. These methods also can be repeated or used in various combinations, e.g., to generate variations in gene/ message expression, message translation or message stability. In another aspect, the genetic composition of a cell is altered by, e.g., modification of a homologous gene *ex vivo*, followed by its reinsertion into the cell.

A nucleic acid of the invention can be altered by any means. For example, random or stochastic methods, or, non-stochastic, or "directed evolution," methods, see, e.g., U.S. Patent No. 6,361,974. Methods for random mutation of genes are well known in the art, see, e.g., U.S. Patent No. 5,830,696. For example, mutagens can be used to randomly mutate a gene. Mutagens include, e.g., ultraviolet light or gamma irradiation, or a chemical mutagen, e.g., mitomycin, nitrous acid, photoactivated psoralens, alone or in combination, to induce DNA breaks amenable to repair by recombination. Other chemical mutagens include, for example, sodium bisulfite, nitrous acid, hydroxylamine, hydrazine or formic acid. Other mutagens are analogues of nucleotide precursors, e.g., nitrosoguanidine, 5-bromouracil, 2-aminopurine, or acridine. These agents can be added to a PCR reaction in place of the nucleotide precursor thereby mutating the sequence. Intercalating agents such as proflavine, acriflavine, quinacrine and the like can also be used.

Any technique in molecular biology can be used, e.g., random PCR mutagenesis, see, e.g., Rice (1992) *Proc. Natl. Acad. Sci. USA* 89:5467-5471; or, combinatorial multiple cassette mutagenesis, see, e.g., Crameri (1995) *Biotechniques* 18:194-196. Alternatively, nucleic acids, e.g., genes, can be reassembled after random, or

5 "stochastic," fragmentation, see, e.g., U.S. Patent Nos. 6,291,242; 6,287,862; 6,287,861; 5,955,358; 5,830,721; 5,824,514; 5,811,238; 5,605,793. In alternative aspects, modifications, additions or deletions are introduced by error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette

10 mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly (e.g., GeneReassembly™, see, e.g., U.S. Patent No. 6,537,776), gene site saturated mutagenesis (GSSM™), synthetic ligation reassembly (SLR), recombination, recursive sequence recombination, phosphothioate-modified DNA

mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis,

15 radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation, and/or a combination of these and other methods.

The following publications describe a variety of recursive recombination procedures and/or methods which can be incorporated into the methods of the invention:

20 Stemmer (1999) "Molecular breeding of viruses for targeting and other clinical properties" *Tumor Targeting* 4:1-4; Ness (1999) *Nature Biotechnology* 17:893-896; Chang (1999) "Evolution of a cytokine using DNA family shuffling" *Nature Biotechnology* 17:793-797; Minshull (1999) "Protein evolution by molecular breeding" *Current Opinion in Chemical Biology* 3:284-290; Christians (1999) "Directed evolution of thymidine kinase for AZT

25 phosphorylation using DNA family shuffling" *Nature Biotechnology* 17:259-264; Crameri (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" *Nature* 391:288-291; Crameri (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling," *Nature Biotechnology* 15:436-438; Zhang (1997) "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and

30 screening" *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Patten et al. (1997) "Applications of DNA Shuffling to Pharmaceuticals and Vaccines" *Current Opinion in Biotechnology* 8:724-733; Crameri et al. (1996) "Construction and evolution of antibody-phage libraries by DNA shuffling" *Nature Medicine* 2:100-103; Gates et al. (1996) "Affinity selective isolation of ligands from peptide libraries through display on a lac repressor 'headpiece dimer'" *Journal*

of Molecular Biology 255:373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" In: The Encyclopedia of Molecular Biology. VCH Publishers, New York. pp.447-457; Crameri and Stemmer (1995) "Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes" *BioTechniques* 18:194-195; Stemmer et al. (1995) "Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides" *Gene*, 164:49-53; Stemmer (1995) "The Evolution of Molecular Computation" *Science* 270: 1510; Stemmer (1995) "Searching Sequence Space" *Bio/Technology* 13:549-553; Stemmer (1994) "Rapid evolution of a protein in vitro by DNA shuffling" *Nature* 370:389-391; and Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution." *Proc. Natl. Acad. Sci. USA* 91:10747-10751.

Mutational methods of generating diversity include, for example, site-directed mutagenesis (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview" *Anal Biochem.* 254(2): 157-178; Dale et al. (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" *Methods Mol. Biol.* 57:369-374; Smith (1985) "In vitro mutagenesis" *Ann. Rev. Genet.* 19:423-462; Botstein & Shortle (1985) "Strategies and applications of in vitro mutagenesis" *Science* 229:1193-1201; Carter (1986) "Site-directed mutagenesis" *Biochem. J.* 237:1-7; and Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in *Nucleic Acids & Molecular Biology* (Eckstein, F. and Lilley, D. M. J. eds., Springer Verlag, Berlin)); mutagenesis using uracil containing templates (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel et al. (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Methods in Enzymol.* 154, 367-382; and Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities" *Science* 242:240-245); oligonucleotide-directed mutagenesis (*Methods in Enzymol.* 100: 468-500 (1983); *Methods in Enzymol.* 154: 329-350 (1987); Zoller (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" *Nucleic Acids Res.* 10:6487-6500; Zoller & Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" *Methods in Enzymol.* 100:468-500; and Zoller (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" *Methods in Enzymol.* 154:329-350); phosphorothioate-modified DNA mutagenesis (Taylor (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" *Nucl. Acids Res.* 13: 8749-8764; Taylor (1985) "The rapid

generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" Nucl. Acids Res. 13: 8765-8787 (1985); Nakamaye (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" Nucl. Acids Res. 14: 9679-9698; Sayers (1988) "Y-T
 5 Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" Nucl. Acids Res. 16:791-802; and Sayers et al. (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" Nucl. Acids Res. 16: 803-814); mutagenesis using gapped duplex DNA (Kramer et al. (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation
 10 construction" Nucl. Acids Res. 12: 9441-9456; Kramer & Fritz (1987) Methods in Enzymol. "Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154:350-367; Kramer (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" Nucl. Acids Res. 16: 7207; and Fritz (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure
 15 without enzymatic reactions *in vitro*" Nucl. Acids Res. 16: 6987-6999).

Additional protocols that can be used to practice the invention include point mismatch repair (Kramer (1984) "Point Mismatch Repair" Cell 38:879-887), mutagenesis using repair-deficient host strains (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" Nucl. Acids Res. 13: 4431-4443; and Carter (1987)
 20 "Improved oligonucleotide-directed mutagenesis using M13 vectors" Methods in Enzymol. 154: 382-403), deletion mutagenesis (Eghtedarzadeh (1986) "Use of oligonucleotides to generate large deletions" Nucl. Acids Res. 14: 5115), restriction-selection and restriction-selection and restriction-purification (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" Phil. Trans. R. Soc. Lond. A 317:
 25 415-423), mutagenesis by total gene synthesis (Nambiar et al. (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" Science 223: 1299-1301; Sakamar and Khorana (1988) "Total synthesis and expression of a gene for the α -subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" Nucl. Acids Res. 14: 6361-6372; Wells et al. (1985) "Cassette mutagenesis: an efficient method for generation of
 30 multiple mutations at defined sites" Gene 34:315-323; and Grundstrom et al. (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" Nucl. Acids Res. 13: 3305-3316), double-strand break repair (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" Current Opinion in Biotechnology 4:450-455.
 "Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a

method for site-specific mutagenesis" Proc. Natl. Acad. Sci. USA, 83:7177-7181). Additional details on many of the above methods can be found in Methods in Enzymology Volume 154, which also describes useful controls for trouble-shooting problems with various mutagenesis methods.

5 Protocols that can be used to practice the invention are described, e.g., in U.S. Patent Nos. 5,605,793 to Stemmer (Feb. 25, 1997), "Methods for In Vitro Recombination;" U.S. Pat. No. 5,811,238 to Stemmer et al. (Sep. 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" U.S. Pat. No. 5,830,721 to Stemmer et al. (Nov. 3, 1998), "DNA Mutagenesis by Random
10 Fragmentation and Reassembly;" U.S. Pat. No. 5,834,252 to Stemmer, et al. (Nov. 10, 1998) "End-Complementary Polymerase Reaction;" U.S. Pat. No. 5,837,458 to Minshull, et al. (Nov. 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering;" WO 95/22625, Stemmer and Crameri, "Mutagenesis by Random Fragmentation and Reassembly;" WO 96/33207 by Stemmer and Lipschutz "End Complementary Polymerase Chain
15 Reaction;" WO 97/20078 by Stemmer and Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination;" WO 97/35966 by Minshull and Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering;" WO 99/41402 by Punnonen et al. "Targeting of Genetic Vaccine Vectors;" WO 99/41383 by Punnonen et al. "Antigen Library Immunization;" WO 99/41369 by
20 Punnonen et al. "Genetic Vaccine Vector Engineering;" WO 99/41368 by Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines;" EP 752008 by Stemmer and Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly;" EP 0932670 by Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination;" WO 99/23107 by Stemmer et al., "Modification of Virus Tropism and Host
25 Range by Viral Genome Shuffling;" WO 99/21979 by Apt et al., "Human Papillomavirus Vectors;" WO 98/31837 by del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination;" WO 98/27230 by Patten and Stemmer, "Methods and Compositions for Polypeptide Engineering;" WO 98/27230 by Stemmer et al., "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection," WO
30 00/00632, "Methods for Generating Highly Diverse Libraries," WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences," WO 98/42832 by Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers," WO 99/29902 by Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences," WO 98/41653 by Vind, "An in Vitro Method for Construction

of a DNA Library," WO 98/41622 by Borchert et al., "Method for Constructing a Library Using DNA Shuffling," and WO 98/42727 by Pati and Zarling, "Sequence Alterations using Homologous Recombination."

Protocols that can be used to practice the invention (providing details
5 regarding various diversity generating methods) are described, e.g., in U.S. Patent application serial no. (USSN) 09/407,800, "SHUFFLING OF CODON ALTERED GENES" by Patten et al. filed Sep. 28, 1999; "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION" by del Cardayre et al., United States Patent No. 6,379,964; "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID
10 RECOMBINATION" by Crameri et al., United States Patent Nos. 6,319,714; 6,368,861; 6,376,246; 6,423,542; 6,426,224 and PCT/US00/01203; "USE OF CODON-VARIED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" by Welch et al., United States Patent No. 6,436,675; "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS"
15 by Selifonov et al., filed Jan. 18, 2000, (PCT/US00/01202) and, e.g. "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" by Selifonov et al., filed Jul. 18, 2000 (U.S. Ser. No. 09/618,579); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" by Selifonov and Stemmer, filed Jan. 18, 2000
20 (PCT/US00/01138); and "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" by Affholter, filed Sep. 6, 2000 (U.S. Ser. No. 09/656,549); and United States Patent Nos. 6,177,263; 6,153,410.

Non-stochastic, or "directed evolution," methods include, e.g., saturation
25 mutagenesis (GSSM™), synthetic ligation reassembly (SLR), or a combination thereof are used to modify the nucleic acids of the invention to generate xylanases with new or altered properties (e.g., activity under highly acidic or alkaline conditions, high or low temperatures, and the like). Polypeptides encoded by the modified nucleic acids can be screened for an activity before testing for xylan hydrolysis or other activity. Any testing modality or protocol
30 can be used, e.g., using a capillary array platform. See, e.g., U.S. Patent Nos. 6,361,974; 6,280,926; 5,939,250.

Saturation mutagenesis, or, GSSM™

In one aspect, codon primers containing a degenerate N,N,G/T sequence are used to introduce point mutations into a polynucleotide, e.g., a xylanase or an antibody of the invention, so as to generate a set of progeny polypeptides in which a full range of single amino acid substitutions is represented at each amino acid position, e.g., an amino acid residue in an enzyme active site or ligand binding site targeted to be modified. These oligonucleotides can comprise a contiguous first homologous sequence, a degenerate N,N,G/T sequence, and, optionally, a second homologous sequence. The downstream progeny translational products from the use of such oligonucleotides include all possible amino acid changes at each amino acid site along the polypeptide, because the degeneracy of the N,N,G/T sequence includes codons for all 20 amino acids. In one aspect, one such degenerate oligonucleotide (comprised of, e.g., one degenerate N,N,G/T cassette) is used for subjecting each original codon in a parental polynucleotide template to a full range of codon substitutions. In another aspect, at least two degenerate cassettes are used – either in the same oligonucleotide or not, for subjecting at least two original codons in a parental polynucleotide template to a full range of codon substitutions. For example, more than one N,N,G/T sequence can be contained in one oligonucleotide to introduce amino acid mutations at more than one site. This plurality of N,N,G/T sequences can be directly contiguous, or separated by one or more additional nucleotide sequence(s). In another aspect, oligonucleotides serviceable for introducing additions and deletions can be used either alone or in combination with the codons containing an N,N,G/T sequence, to introduce any combination or permutation of amino acid additions, deletions, and/or substitutions.

In one aspect, simultaneous mutagenesis of two or more contiguous amino acid positions is done using an oligonucleotide that contains contiguous N,N,G/T triplets, i.e. a degenerate (N,N,G/T)_n sequence. In another aspect, degenerate cassettes having less degeneracy than the N,N,G/T sequence are used. For example, it may be desirable in some instances to use (e.g. in an oligonucleotide) a degenerate triplet sequence comprised of only one N, where said N can be in the first second or third position of the triplet. Any other bases including any combinations and permutations thereof can be used in the remaining two positions of the triplet. Alternatively, it may be desirable in some instances to use (e.g. in an oligo) a degenerate N,N,N triplet sequence.

In one aspect, use of degenerate triplets (e.g., N,N,G/T triplets) allows for systematic and easy generation of a full range of possible natural amino acids (for a total of 20 amino acids) into each and every amino acid position in a polypeptide (in alternative

aspects, the methods also include generation of less than all possible substitutions per amino acid residue, or codon, position). For example, for a 100 amino acid polypeptide, 2000 distinct species (i.e. 20 possible amino acids per position X 100 amino acid positions) can be generated. Through the use of an oligonucleotide or set of oligonucleotides containing a degenerate N,N,G/T triplet, 32 individual sequences can code for all 20 possible natural amino acids. Thus, in a reaction vessel in which a parental polynucleotide sequence is subjected to saturation mutagenesis using at least one such oligonucleotide, there are generated 32 distinct progeny polynucleotides encoding 20 distinct polypeptides. In contrast, the use of a non-degenerate oligonucleotide in site-directed mutagenesis leads to only one progeny polypeptide product per reaction vessel. Nondegenerate oligonucleotides can optionally be used in combination with degenerate primers disclosed; for example, nondegenerate oligonucleotides can be used to generate specific point mutations in a working polynucleotide. This provides one means to generate specific silent point mutations, point mutations leading to corresponding amino acid changes, and point mutations that cause the generation of stop codons and the corresponding expression of polypeptide fragments.

In one aspect, each saturation mutagenesis reaction vessel contains polynucleotides encoding at least 20 progeny polypeptide (e.g., xylanases) molecules such that all 20 natural amino acids are represented at the one specific amino acid position corresponding to the codon position mutagenized in the parental polynucleotide (other aspects use less than all 20 natural combinations). The 32-fold degenerate progeny polypeptides generated from each saturation mutagenesis reaction vessel can be subjected to clonal amplification (e.g. cloned into a suitable host, e.g., *E. coli* host, using, e.g., an expression vector) and subjected to expression screening. When an individual progeny polypeptide is identified by screening to display a favorable change in property (when compared to the parental polypeptide, such as increased xylan hydrolysis activity under alkaline or acidic conditions), it can be sequenced to identify the correspondingly favorable amino acid substitution contained therein.

In one aspect, upon mutagenizing each and every amino acid position in a parental polypeptide using saturation mutagenesis as disclosed herein, favorable amino acid changes may be identified at more than one amino acid position. One or more new progeny molecules can be generated that contain a combination of all or part of these favorable amino acid substitutions. For example, if 2 specific favorable amino acid changes are identified in each of 3 amino acid positions in a polypeptide, the permutations include 3 possibilities at each position (no change from the original amino acid, and each of two favorable changes)

and 3 positions. Thus, there are $3 \times 3 \times 3$ or 27 total possibilities, including 7 that were previously examined - 6 single point mutations (i.e. 2 at each of three positions) and no change at any position.

5 In yet another aspect, site-saturation mutagenesis can be used together with shuffling, chimerization, recombination and other mutagenizing processes, along with screening. This invention provides for the use of any mutagenizing process(es), including saturation mutagenesis, in an iterative manner. In one exemplification, the iterative use of any mutagenizing process(es) is used in combination with screening.

The invention also provides for the use of proprietary codon primers
10 (containing a degenerate N,N,N sequence) to introduce point mutations into a polynucleotide, so as to generate a set of progeny polypeptides in which a full range of single amino acid substitutions is represented at each amino acid position (gene site saturated mutagenesis (GSSM™)). The oligos used are comprised contiguously of a first homologous sequence, a degenerate N,N,N sequence and preferably but not necessarily a second homologous
15 sequence. The downstream progeny translational products from the use of such oligos include all possible amino acid changes at each amino acid site along the polypeptide, because the degeneracy of the N,N,N sequence includes codons for all 20 amino acids.

In one aspect, one such degenerate oligo (comprised of one degenerate N,N,N cassette) is used for subjecting each original codon in a parental polynucleotide template to a
20 full range of codon substitutions. In another aspect, at least two degenerate N,N,N cassettes are used – either in the same oligo or not, for subjecting at least two original codons in a parental polynucleotide template to a full range of codon substitutions. Thus, more than one N,N,N sequence can be contained in one oligo to introduce amino acid mutations at more than one site. This plurality of N,N,N sequences can be directly contiguous, or separated by
25 one or more additional nucleotide sequence(s). In another aspect, oligos serviceable for introducing additions and deletions can be used either alone or in combination with the codons containing an N,N,N sequence, to introduce any combination or permutation of amino acid additions, deletions and/or substitutions.

In a particular exemplification, it is possible to simultaneously mutagenize two
30 or more contiguous amino acid positions using an oligo that contains contiguous N,N,N triplets, i.e. a degenerate $(N,N,N)_n$ sequence.

In another aspect, the present invention provides for the use of degenerate cassettes having less degeneracy than the N,N,N sequence. For example, it may be desirable in some instances to use (e.g. in an oligo) a degenerate triplet sequence comprised of only one

N, where the N can be in the first second or third position of the triplet. Any other bases including any combinations and permutations thereof can be used in the remaining two positions of the triplet. Alternatively, it may be desirable in some instances to use (*e.g.*, in an oligo) a degenerate N,N,N triplet sequence, N,N,G/T, or an N,N, G/C triplet sequence.

5 It is appreciated, however, that the use of a degenerate triplet (such as N,N,G/T or an N,N, G/C triplet sequence) as disclosed in the instant invention is advantageous for several reasons. In one aspect, this invention provides a means to systematically and fairly easily generate the substitution of the full range of possible amino acids (for a total of 20 amino acids) into each and every amino acid position in a polypeptide. 10 Thus, for a 100 amino acid polypeptide, the invention provides a way to systematically and fairly easily generate 2000 distinct species (*i.e.*, 20 possible amino acids per position times 100 amino acid positions). It is appreciated that there is provided, through the use of an oligo containing a degenerate N,N,G/T or an N,N, G/C triplet sequence, 32 individual sequences that code for 20 possible amino acids. Thus, in a reaction vessel in which a parental 15 polynucleotide sequence is subjected to saturation mutagenesis using one such oligo, there are generated 32 distinct progeny polynucleotides encoding 20 distinct polypeptides. In contrast, the use of a non-degenerate oligo in site-directed mutagenesis leads to only one progeny polypeptide product per reaction vessel.

20 This invention also provides for the use of nondegenerate oligos, which can optionally be used in combination with degenerate primers disclosed. It is appreciated that in some situations, it is advantageous to use nondegenerate oligos to generate specific point mutations in a working polynucleotide. This provides a means to generate specific silent point mutations, point mutations leading to corresponding amino acid changes and point mutations that cause the generation of stop codons and the corresponding expression of 25 polypeptide fragments.

Thus, in one aspect of this invention, each saturation mutagenesis reaction vessel contains polynucleotides encoding at least 20 progeny polypeptide molecules such that all 20 amino acids are represented at the one specific amino acid position corresponding to the codon position mutagenized in the parental polynucleotide. The 32-fold degenerate 30 progeny polypeptides generated from each saturation mutagenesis reaction vessel can be subjected to clonal amplification (*e.g.*, cloned into a suitable *E. coli* host using an expression vector) and subjected to expression screening. When an individual progeny polypeptide is identified by screening to display a favorable change in property (when compared to the

parental polypeptide), it can be sequenced to identify the correspondingly favorable amino acid substitution contained therein.

It is appreciated that upon mutagenizing each and every amino acid position in a parental polypeptide using saturation mutagenesis as disclosed herein, favorable amino acid changes may be identified at more than one amino acid position. One or more new progeny molecules can be generated that contain a combination of all or part of these favorable amino acid substitutions. For example, if 2 specific favorable amino acid changes are identified in each of 3 amino acid positions in a polypeptide, the permutations include 3 possibilities at each position (no change from the original amino acid and each of two favorable changes) and 3 positions. Thus, there are $3 \times 3 \times 3$ or 27 total possibilities, including 7 that were previously examined - 6 single point mutations (*i.e.*, 2 at each of three positions) and no change at any position.

Thus, in a non-limiting exemplification, this invention provides for the use of saturation mutagenesis in combination with additional mutagenization processes, such as process where two or more related polynucleotides are introduced into a suitable host cell such that a hybrid polynucleotide is generated by recombination and reductive reassortment.

In addition to performing mutagenesis along the entire sequence of a gene, the instant invention provides that mutagenesis can be used to replace each of any number of bases in a polynucleotide sequence, wherein the number of bases to be mutagenized is preferably every integer from 15 to 100,000. Thus, instead of mutagenizing every position along a molecule, one can subject every or a discrete number of bases (preferably a subset totaling from 15 to 100,000) to mutagenesis. Preferably, a separate nucleotide is used for mutagenizing each position or group of positions along a polynucleotide sequence. A group of 3 positions to be mutagenized may be a codon. The mutations are preferably introduced using a mutagenic primer, containing a heterologous cassette, also referred to as a mutagenic cassette. Exemplary cassettes can have from 1 to 500 bases. Each nucleotide position in such heterologous cassettes be N, A, C, G, T, A/C, A/G, A/T, C/G, C/T, G/T, C/G/T, A/G/T, A/C/T, A/C/G, or E, where E is any base that is not A, C, G, or T (E can be referred to as a designer oligo).

In a general sense, saturation mutagenesis is comprised of mutagenizing a complete set of mutagenic cassettes (wherein each cassette is preferably about 1-500 bases in length) in defined polynucleotide sequence to be mutagenized (wherein the sequence to be mutagenized is preferably from about 15 to 100,000 bases in length). Thus, a group of mutations (ranging from 1 to 100 mutations) is introduced into each cassette to be

mutagenized. A grouping of mutations to be introduced into one cassette can be different or the same from a second grouping of mutations to be introduced into a second cassette during the application of one round of saturation mutagenesis. Such groupings are exemplified by deletions, additions, groupings of particular codons and groupings of particular nucleotide cassettes.

Defined sequences to be mutagenized include a whole gene, pathway, cDNA, an entire open reading frame (ORF) and entire promoter, enhancer, repressor/transactivator, origin of replication, intron, operator, or any polynucleotide functional group. Generally, a "defined sequences" for this purpose may be any polynucleotide that a 15 base-polynucleotide sequence and polynucleotide sequences of lengths between 15 bases and 15,000 bases (this invention specifically names every integer in between). Considerations in choosing groupings of codons include types of amino acids encoded by a degenerate mutagenic cassette.

In one exemplification a grouping of mutations that can be introduced into a mutagenic cassette, this invention specifically provides for degenerate codon substitutions (using degenerate oligos) that code for 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20 amino acids at each position and a library of polypeptides encoded thereby.

Synthetic Ligation Reassembly (SLR)

The invention provides a non-stochastic gene modification system termed "synthetic ligation reassembly," or simply "SLR," a "directed evolution process," to generate polypeptides, e.g., xylanases or antibodies of the invention, with new or altered properties.

SLR is a method of ligating oligonucleotide fragments together non-stochastically. This method differs from stochastic oligonucleotide shuffling in that the nucleic acid building blocks are not shuffled, concatenated or chimerized randomly, but rather are assembled non-stochastically. See, e.g., U.S. Patent Application Serial No. (USSN) 09/332,835 entitled "Synthetic Ligation Reassembly in Directed Evolution" and filed on June 14, 1999 ("USSN 09/332,835"). In one aspect, SLR comprises the following steps: (a) providing a template polynucleotide, wherein the template polynucleotide comprises sequence encoding a homologous gene; (b) providing a plurality of building block polynucleotides, wherein the building block polynucleotides are designed to cross-over reassemble with the template polynucleotide at a predetermined sequence, and a building block polynucleotide comprises a sequence that is a variant of the homologous gene and a sequence homologous to the template polynucleotide flanking the variant sequence; (c)

combining a building block polynucleotide with a template polynucleotide such that the building block polynucleotide cross-over reassembles with the template polynucleotide to generate polynucleotides comprising homologous gene sequence variations.

SLR does not depend on the presence of high levels of homology between polynucleotides to be rearranged. Thus, this method can be used to non-stochastically generate libraries (or sets) of progeny molecules comprised of over 10^{100} different chimeras. SLR can be used to generate libraries comprised of over 10^{1000} different progeny chimeras. Thus, aspects of the present invention include non-stochastic methods of producing a set of finalized chimeric nucleic acid molecule having an overall assembly order that is chosen by design. This method includes the steps of generating by design a plurality of specific nucleic acid building blocks having serviceable mutually compatible ligatable ends, and assembling these nucleic acid building blocks, such that a designed overall assembly order is achieved.

The mutually compatible ligatable ends of the nucleic acid building blocks to be assembled are considered to be "serviceable" for this type of ordered assembly if they enable the building blocks to be coupled in predetermined orders. Thus, the overall assembly order in which the nucleic acid building blocks can be coupled is specified by the design of the ligatable ends. If more than one assembly step is to be used, then the overall assembly order in which the nucleic acid building blocks can be coupled is also specified by the sequential order of the assembly step(s). In one aspect, the annealed building pieces are treated with an enzyme, such as a ligase (e.g. T4 DNA ligase), to achieve covalent bonding of the building pieces.

In one aspect, the design of the oligonucleotide building blocks is obtained by analyzing a set of progenitor nucleic acid sequence templates that serve as a basis for producing a progeny set of finalized chimeric polynucleotides. These parental oligonucleotide templates thus serve as a source of sequence information that aids in the design of the nucleic acid building blocks that are to be mutagenized, e.g., chimerized or shuffled. In one aspect of this method, the sequences of a plurality of parental nucleic acid templates are aligned in order to select one or more demarcation points. The demarcation points can be located at an area of homology, and are comprised of one or more nucleotides. These demarcation points are preferably shared by at least two of the progenitor templates. The demarcation points can thereby be used to delineate the boundaries of oligonucleotide building blocks to be generated in order to rearrange the parental polynucleotides. The demarcation points identified and selected in the progenitor molecules serve as potential chimerization points in the assembly of the final chimeric progeny molecules. A demarcation

point can be an area of homology (comprised of at least one homologous nucleotide base) shared by at least two parental polynucleotide sequences. Alternatively, a demarcation point can be an area of homology that is shared by at least half of the parental polynucleotide sequences, or, it can be an area of homology that is shared by at least two thirds of the parental polynucleotide sequences. Even more preferably a serviceable demarcation points is an area of homology that is shared by at least three fourths of the parental polynucleotide sequences, or, it can be shared by at almost all of the parental polynucleotide sequences. In one aspect, a demarcation point is an area of homology that is shared by all of the parental polynucleotide sequences.

In one aspect, a ligation reassembly process is performed exhaustively in order to generate an exhaustive library of progeny chimeric polynucleotides. In other words, all possible ordered combinations of the nucleic acid building blocks are represented in the set of finalized chimeric nucleic acid molecules. At the same time, in another aspect, the assembly order (i.e. the order of assembly of each building block in the 5' to 3' sequence of each finalized chimeric nucleic acid) in each combination is by design (or non-stochastic) as described above. Because of the non-stochastic nature of this invention, the possibility of unwanted side products is greatly reduced.

In another aspect, the ligation reassembly method is performed systematically. For example, the method is performed in order to generate a systematically compartmentalized library of progeny molecules, with compartments that can be screened systematically, e.g. one by one. In other words this invention provides that, through the selective and judicious use of specific nucleic acid building blocks, coupled with the selective and judicious use of sequentially stepped assembly reactions, a design can be achieved where specific sets of progeny products are made in each of several reaction vessels. This allows a systematic examination and screening procedure to be performed. Thus, these methods allow a potentially very large number of progeny molecules to be examined systematically in smaller groups. Because of its ability to perform chimerizations in a manner that is highly flexible yet exhaustive and systematic as well, particularly when there is a low level of homology among the progenitor molecules, these methods provide for the generation of a library (or set) comprised of a large number of progeny molecules. Because of the non-stochastic nature of the instant ligation reassembly invention, the progeny molecules generated preferably comprise a library of finalized chimeric nucleic acid molecules having an overall assembly order that is chosen by design. The saturation mutagenesis and optimized directed evolution methods also can be used to generate different progeny

molecular species. It is appreciated that the invention provides freedom of choice and control regarding the selection of demarcation points, the size and number of the nucleic acid building blocks, and the size and design of the couplings. It is appreciated, furthermore, that the requirement for intermolecular homology is highly relaxed for the operability of this invention. In fact, demarcation points can even be chosen in areas of little or no intermolecular homology. For example, because of codon wobble, i.e. the degeneracy of codons, nucleotide substitutions can be introduced into nucleic acid building blocks without altering the amino acid originally encoded in the corresponding progenitor template. Alternatively, a codon can be altered such that the coding for an originally amino acid is altered. This invention provides that such substitutions can be introduced into the nucleic acid building block in order to increase the incidence of intermolecular homologous demarcation points and thus to allow an increased number of couplings to be achieved among the building blocks, which in turn allows a greater number of progeny chimeric molecules to be generated.

Synthetic gene reassembly

In one aspect, the present invention provides a non-stochastic method termed synthetic gene reassembly (e.g., GeneReassembly™, see, e.g., U.S. Patent No. 6,537,776), which differs from stochastic shuffling in that the nucleic acid building blocks are not shuffled or concatenated or chimerized randomly, but rather are assembled non-stochastically.

The synthetic gene reassembly method does not depend on the presence of a high level of homology between polynucleotides to be shuffled. The invention can be used to non-stochastically generate libraries (or sets) of progeny molecules comprised of over 10^{100} different chimeras. Conceivably, synthetic gene reassembly can even be used to generate libraries comprised of over 10^{1000} different progeny chimeras.

Thus, in one aspect, the invention provides a non-stochastic method of producing a set of finalized chimeric nucleic acid molecules having an overall assembly order that is chosen by design, which method is comprised of the steps of generating by design a plurality of specific nucleic acid building blocks having serviceable mutually compatible ligatable ends and assembling these nucleic acid building blocks, such that a designed overall assembly order is achieved.

In one aspect, synthetic gene reassembly comprises a method of: 1) preparing a progeny generation of molecule(s) (including a molecule comprising a polynucleotide

sequence, e.g., a molecule comprising a polypeptide coding sequence), that is mutagenized to achieve at least one point mutation, addition, deletion, &/or chimerization, from one or more ancestral or parental generation template(s); 2) screening the progeny generation molecule(s), e.g., using a high throughput method, for at least one property of interest (such as an improvement in an enzyme activity); 3) optionally obtaining &/or cataloguing structural &/or and functional information regarding the parental &/or progeny generation molecules; and 4) optionally repeating any of steps 1) to 3). In one aspect, there is generated (e.g., from a parent polynucleotide template), in what is termed "codon site-saturation mutagenesis," a progeny generation of polynucleotides, each having at least one set of up to three contiguous point mutations (i.e. different bases comprising a new codon), such that every codon (or every family of degenerate codons encoding the same amino acid) is represented at each codon position. Corresponding to, and encoded by, this progeny generation of polynucleotides, there is also generated a set of progeny polypeptides, each having at least one single amino acid point mutation. In a one aspect, there is generated, in what is termed "amino acid site-saturation mutagenesis", one such mutant polypeptide for each of the 19 naturally encoded polypeptide-forming alpha-amino acid substitutions at each and every amino acid position along the polypeptide. This yields, for each and every amino acid position along the parental polypeptide, a total of 20 distinct progeny polypeptides including the original amino acid, or potentially more than 21 distinct progeny polypeptides if additional amino acids are used either instead of or in addition to the 20 naturally encoded amino acids

Thus, in another aspect, this approach is also serviceable for generating mutants containing, in addition to &/or in combination with the 20 naturally encoded polypeptide-forming alpha-amino acids, other rare &/or not naturally-encoded amino acids and amino acid derivatives. In yet another aspect, this approach is also serviceable for generating mutants by the use of, in addition to &/or in combination with natural or unaltered codon recognition systems of suitable hosts, altered, mutagenized, &/or designer codon recognition systems (such as in a host cell with one or more altered tRNA molecules).

In yet another aspect, this invention relates to recombination and more specifically to a method for preparing polynucleotides encoding a polypeptide by a method of *in vivo* re-assortment of polynucleotide sequences containing regions of partial homology, assembling the polynucleotides to form at least one polynucleotide and screening the polynucleotides for the production of polypeptide(s) having a useful property.

In yet another aspect, this invention is serviceable for analyzing and cataloguing, with respect to any molecular property (e.g. an enzymatic activity) or combination of properties allowed by current technology, the effects of any mutational change achieved (including particularly saturation mutagenesis). Thus, a comprehensive
5 method is provided for determining the effect of changing each amino acid in a parental polypeptide into each of at least 19 possible substitutions. This allows each amino acid in a parental polypeptide to be characterized and catalogued according to its spectrum of potential effects on a measurable property of the polypeptide.

In one aspect, an intron may be introduced into a chimeric progeny molecule
10 by way of a nucleic acid building block. Introns often have consensus sequences at both termini in order to render them operational. In addition to enabling gene splicing, introns may serve an additional purpose by providing sites of homology to other nucleic acids to enable homologous recombination. For this purpose, and potentially others, it may be sometimes desirable to generate a large nucleic acid building block for introducing an intron.
15 If the size is overly large easily generating by direct chemical synthesis of two single stranded oligos, such a specialized nucleic acid building block may also be generated by direct chemical synthesis of more than two single stranded oligos or by using a polymerase-based amplification reaction

The mutually compatible ligatable ends of the nucleic acid building blocks to
20 be assembled are considered to be "serviceable" for this type of ordered assembly if they enable the building blocks to be coupled in predetermined orders. Thus, in one aspect, the overall assembly order in which the nucleic acid building blocks can be coupled is specified by the design of the ligatable ends and, if more than one assembly step is to be used, then the overall assembly order in which the nucleic acid building blocks can be coupled is also
25 specified by the sequential order of the assembly step(s). In a one aspect of the invention, the annealed building pieces are treated with an enzyme, such as a ligase (e.g., T4 DNA ligase) to achieve covalent bonding of the building pieces.

Coupling can occur in a manner that does not make use of every nucleotide in a participating overhang. The coupling is particularly lively to survive (e.g. in a transformed
30 host) if the coupling reinforced by treatment with a ligase enzyme to form what may be referred to as a "gap ligation" or a "gapped ligation". This type of coupling can contribute to generation of unwanted background product(s), but it can also be used advantageously increase the diversity of the progeny library generated by the designed ligation reassembly. Certain overhangs are able to undergo self-coupling to form a palindromic coupling. A

coupling is strengthened substantially if it is reinforced by treatment with a ligase enzyme. Lack of 5' phosphates on these overhangs can be used advantageously to prevent this type of palindromic self-ligation. Accordingly, this invention provides that nucleic acid building blocks can be chemically made (or ordered) that lack a 5' phosphate group. Alternatively, 5 they can be removed, e.g. by treatment with a phosphatase enzyme, such as a calf intestinal alkaline phosphatase (CIAP), in order to prevent palindromic self-ligations in ligation reassembly processes.

In another aspect, the design of nucleic acid building blocks is obtained upon analysis of the sequences of a set of progenitor nucleic acid templates that serve as a basis for 10 producing a progeny set of finalized chimeric nucleic acid molecules. These progenitor nucleic acid templates thus serve as a source of sequence information that aids in the design of the nucleic acid building blocks that are to be mutagenized, *i.e.* chimerized or shuffled.

In one exemplification, the invention provides for the chimerization of a family of related genes and their encoded family of related products. In a particular 15 exemplification, the encoded products are enzymes. The xylanases of the present invention can be mutagenized in accordance with the methods described herein.

Thus according to one aspect of the invention, the sequences of a plurality of progenitor nucleic acid templates (*e.g.*, polynucleotides of Group A nucleic acid sequences) are aligned in order to select one or more demarcation points, which demarcation points can 20 be located at an area of homology. The demarcation points can be used to delineate the boundaries of nucleic acid building blocks to be generated. Thus, the demarcation points identified and selected in the progenitor molecules serve as potential chimerization points in the assembly of the progeny molecules.

Typically a serviceable demarcation point is an area of homology (comprised 25 of at least one homologous nucleotide base) shared by at least two progenitor templates, but the demarcation point can be an area of homology that is shared by at least half of the progenitor templates, at least two thirds of the progenitor templates, at least three fourths of the progenitor templates and preferably at almost all of the progenitor templates. Even more preferably still a serviceable demarcation point is an area of homology that is shared by all of 30 the progenitor templates.

In one aspect, the gene reassembly process is performed exhaustively in order to generate an exhaustive library. In other words, all possible ordered combinations of the nucleic acid building blocks are represented in the set of finalized chimeric nucleic acid molecules. At the same time, the assembly order (*i.e.* the order of assembly of each building

block in the 5' to 3' sequence of each finalized chimeric nucleic acid) in each combination is by design (or non-stochastic). Because of the non-stochastic nature of the method, the possibility of unwanted side products is greatly reduced.

In another aspect, the method provides that the gene reassembly process is performed systematically, for example to generate a systematically compartmentalized library, with compartments that can be screened systematically, *e.g.*, one by one. In other words the invention provides that, through the selective and judicious use of specific nucleic acid building blocks, coupled with the selective and judicious use of sequentially stepped assembly reactions, an experimental design can be achieved where specific sets of progeny products are made in each of several reaction vessels. This allows a systematic examination and screening procedure to be performed. Thus, it allows a potentially very large number of progeny molecules to be examined systematically in smaller groups.

Because of its ability to perform chimerizations in a manner that is highly flexible yet exhaustive and systematic as well, particularly when there is a low level of homology among the progenitor molecules, the instant invention provides for the generation of a library (or set) comprised of a large number of progeny molecules. Because of the non-stochastic nature of the instant gene reassembly invention, the progeny molecules generated preferably comprise a library of finalized chimeric nucleic acid molecules having an overall assembly order that is chosen by design. In a particularly aspect, such a generated library is comprised of greater than 10^3 to greater than 10^{1000} different progeny molecular species.

In one aspect, a set of finalized chimeric nucleic acid molecules, produced as described is comprised of a polynucleotide encoding a polypeptide. According to one aspect, this polynucleotide is a gene, which may be a man-made gene. According to another aspect, this polynucleotide is a gene pathway, which may be a man-made gene pathway. The invention provides that one or more man-made genes generated by the invention may be incorporated into a man-made gene pathway, such as pathway operable in a eukaryotic organism (including a plant).

In another exemplification, the synthetic nature of the step in which the building blocks are generated allows the design and introduction of nucleotides (*e.g.*, one or more nucleotides, which may be, for example, codons or introns or regulatory sequences) that can later be optionally removed in an *in vitro* process (*e.g.*, by mutagenesis) or in an *in vivo* process (*e.g.*, by utilizing the gene splicing ability of a host organism). It is appreciated that in many instances the introduction of these nucleotides may also be desirable for many other reasons in addition to the potential benefit of creating a serviceable demarcation point.

Thus, according to another aspect, the invention provides that a nucleic acid building block can be used to introduce an intron. Thus, the invention provides that functional introns may be introduced into a man-made gene of the invention. The invention also provides that functional introns may be introduced into a man-made gene pathway of the invention. Accordingly, the invention provides for the generation of a chimeric polynucleotide that is a man-made gene containing one (or more) artificially introduced intron(s).

Accordingly, the invention also provides for the generation of a chimeric polynucleotide that is a man-made gene pathway containing one (or more) artificially introduced intron(s). Preferably, the artificially introduced intron(s) are functional in one or more host cells for gene splicing much in the way that naturally-occurring introns serve functionally in gene splicing. The invention provides a process of producing man-made intron-containing polynucleotides to be introduced into host organisms for recombination and/or splicing.

A man-made gene produced using the invention can also serve as a substrate for recombination with another nucleic acid. Likewise, a man-made gene pathway produced using the invention can also serve as a substrate for recombination with another nucleic acid. In one aspect, the recombination is facilitated by, or occurs at, areas of homology between the man-made, intron-containing gene and a nucleic acid, which serves as a recombination partner. In one aspect, the recombination partner may also be a nucleic acid generated by the invention, including a man-made gene or a man-made gene pathway. Recombination may be facilitated by or may occur at areas of homology that exist at the one (or more) artificially introduced intron(s) in the man-made gene.

The synthetic gene reassembly method of the invention utilizes a plurality of nucleic acid building blocks, each of which preferably has two ligatable ends. The two ligatable ends on each nucleic acid building block may be two blunt ends (*i.e.* each having an overhang of zero nucleotides), or preferably one blunt end and one overhang, or more preferably still two overhangs.

A useful overhang for this purpose may be a 3' overhang or a 5' overhang. Thus, a nucleic acid building block may have a 3' overhang or alternatively a 5' overhang or alternatively two 3' overhangs or alternatively two 5' overhangs. The overall order in which the nucleic acid building blocks are assembled to form a finalized chimeric nucleic acid molecule is determined by purposeful experimental design and is not random.

In one aspect, a nucleic acid building block is generated by chemical synthesis of two single-stranded nucleic acids (also referred to as single-stranded oligos) and contacting them so as to allow them to anneal to form a double-stranded nucleic acid building block.

A double-stranded nucleic acid building block can be of variable size. The
5 sizes of these building blocks can be small or large. Exemplary sizes for building block range from 1 base pair (not including any overhangs) to 100,000 base pairs (not including any overhangs). Other exemplary size ranges are also provided, which have lower limits of from 1 bp to 10,000 bp (including every integer value in between) and upper limits of from 2 bp to 100,000 bp (including every integer value in between).

10 Many methods exist by which a double-stranded nucleic acid building block can be generated that is serviceable for the invention; and these are known in the art and can be readily performed by the skilled artisan.

According to one aspect, a double-stranded nucleic acid building block is generated by first generating two single stranded nucleic acids and allowing them to anneal to
15 form a double-stranded nucleic acid building block. The two strands of a double-stranded nucleic acid building block may be complementary at every nucleotide apart from any that form an overhang; thus containing no mismatches, apart from any overhang(s). According to another aspect, the two strands of a double-stranded nucleic acid building block are complementary at fewer than every nucleotide apart from any that form an overhang. Thus,
20 according to this aspect, a double-stranded nucleic acid building block can be used to introduce codon degeneracy. Preferably the codon degeneracy is introduced using the site-saturation mutagenesis described herein, using one or more N,N,G/T cassettes or alternatively using one or more N,N,N cassettes.

The *in vivo* recombination method of the invention can be performed blindly
25 on a pool of unknown hybrids or alleles of a specific polynucleotide or sequence. However, it is not necessary to know the actual DNA or RNA sequence of the specific polynucleotide.

The approach of using recombination within a mixed population of genes can be useful for the generation of any useful proteins, for example, interleukin I, antibodies, tPA and growth hormone. This approach may be used to generate proteins having altered
30 specificity or activity. The approach may also be useful for the generation of hybrid nucleic acid sequences, for example, promoter regions, introns, exons, enhancer sequences, 3' untranslated regions or 5' untranslated regions of genes. Thus this approach may be used to generate genes having increased rates of expression. This approach may also be useful in the

study of repetitive DNA sequences. Finally, this approach may be useful to mutate ribozymes or aptamers.

In one aspect the invention described herein is directed to the use of repeated cycles of reductive reassortment, recombination and selection which allow for the directed molecular evolution of highly complex linear sequences, such as DNA, RNA or proteins
5 thorough recombination.

Optimized Directed Evolution System

The invention provides a non-stochastic gene modification system termed “optimized directed evolution system” to generate polypeptides, e.g., xylanases or antibodies
10 of the invention, with new or altered properties. Optimized directed evolution is directed to the use of repeated cycles of reductive reassortment, recombination and selection that allow for the directed molecular evolution of nucleic acids through recombination. Optimized directed evolution allows generation of a large population of evolved chimeric sequences, wherein the generated population is significantly enriched for sequences that have a
15 predetermined number of crossover events.

A crossover event is a point in a chimeric sequence where a shift in sequence occurs from one parental variant to another parental variant. Such a point is normally at the juncture of where oligonucleotides from two parents are ligated together to form a single sequence. This method allows calculation of the correct concentrations of oligonucleotide
20 sequences so that the final chimeric population of sequences is enriched for the chosen number of crossover events. This provides more control over choosing chimeric variants having a predetermined number of crossover events.

In addition, this method provides a convenient means for exploring a tremendous amount of the possible protein variant space in comparison to other systems.
25 Previously, if one generated, for example, 10^{13} chimeric molecules during a reaction, it would be extremely difficult to test such a high number of chimeric variants for a particular activity. Moreover, a significant portion of the progeny population would have a very high number of crossover events which resulted in proteins that were less likely to have increased levels of a particular activity. By using these methods, the population of chimerics molecules can be
30 enriched for those variants that have a particular number of crossover events. Thus, although one can still generate 10^{13} chimeric molecules during a reaction, each of the molecules chosen for further analysis most likely has, for example, only three crossover events. Because the resulting progeny population can be skewed to have a predetermined number of

crossover events, the boundaries on the functional variety between the chimeric molecules is reduced. This provides a more manageable number of variables when calculating which oligonucleotide from the original parental polynucleotides might be responsible for affecting a particular trait.

5 One method for creating a chimeric progeny polynucleotide sequence is to create oligonucleotides corresponding to fragments or portions of each parental sequence. Each oligonucleotide preferably includes a unique region of overlap so that mixing the oligonucleotides together results in a new variant that has each oligonucleotide fragment assembled in the correct order. Additional information can also be found, e.g., in USSN
10 09/332,835; U.S. Patent No. 6,361,974.

 The number of oligonucleotides generated for each parental variant bears a relationship to the total number of resulting crossovers in the chimeric molecule that is ultimately created. For example, three parental nucleotide sequence variants might be provided to undergo a ligation reaction in order to find a chimeric variant having, for
15 example, greater activity at high temperature. As one example, a set of 50 oligonucleotide sequences can be generated corresponding to each portions of each parental variant. Accordingly, during the ligation reassembly process there could be up to 50 crossover events within each of the chimeric sequences. The probability that each of the generated chimeric polynucleotides will contain oligonucleotides from each parental variant in alternating order
20 is very low. If each oligonucleotide fragment is present in the ligation reaction in the same molar quantity it is likely that in some positions oligonucleotides from the same parental polynucleotide will ligate next to one another and thus not result in a crossover event. If the concentration of each oligonucleotide from each parent is kept constant during any ligation step in this example, there is a 1/3 chance (assuming 3 parents) that an oligonucleotide from
25 the same parental variant will ligate within the chimeric sequence and produce no crossover.

 Accordingly, a probability density function (PDF) can be determined to predict the population of crossover events that are likely to occur during each step in a ligation reaction given a set number of parental variants, a number of oligonucleotides corresponding to each variant, and the concentrations of each variant during each step in the
30 ligation reaction. The statistics and mathematics behind determining the PDF is described below. By utilizing these methods, one can calculate such a probability density function, and thus enrich the chimeric progeny population for a predetermined number of crossover events resulting from a particular ligation reaction. Moreover, a target number of crossover events can be predetermined, and the system then programmed to calculate the starting quantities of

each parental oligonucleotide during each step in the ligation reaction to result in a probability density function that centers on the predetermined number of crossover events. These methods are directed to the use of repeated cycles of reductive reassortment, recombination and selection that allow for the directed molecular evolution of a nucleic acid encoding a polypeptide through recombination. This system allows generation of a large population of evolved chimeric sequences, wherein the generated population is significantly enriched for sequences that have a predetermined number of crossover events. A crossover event is a point in a chimeric sequence where a shift in sequence occurs from one parental variant to another parental variant. Such a point is normally at the juncture of where oligonucleotides from two parents are ligated together to form a single sequence. The method allows calculation of the correct concentrations of oligonucleotide sequences so that the final chimeric population of sequences is enriched for the chosen number of crossover events. This provides more control over choosing chimeric variants having a predetermined number of crossover events.

In addition, these methods provide a convenient means for exploring a tremendous amount of the possible protein variant space in comparison to other systems. By using the methods described herein, the population of chimeric molecules can be enriched for those variants that have a particular number of crossover events. Thus, although one can still generate 10^{13} chimeric molecules during a reaction, each of the molecules chosen for further analysis most likely has, for example, only three crossover events. Because the resulting progeny population can be skewed to have a predetermined number of crossover events, the boundaries on the functional variety between the chimeric molecules is reduced. This provides a more manageable number of variables when calculating which oligonucleotide from the original parental polynucleotides might be responsible for affecting a particular trait.

In one aspect, the method creates a chimeric progeny polynucleotide sequence by creating oligonucleotides corresponding to fragments or portions of each parental sequence. Each oligonucleotide preferably includes a unique region of overlap so that mixing the oligonucleotides together results in a new variant that has each oligonucleotide fragment assembled in the correct order. See also USSN 09/332,835.

Determining Crossover Events

Aspects of the invention include a system and software that receive a desired crossover probability density function (PDF), the number of parent genes to be reassembled,

and the number of fragments in the reassembly as inputs. The output of this program is a “fragment PDF” that can be used to determine a recipe for producing reassembled genes, and the estimated crossover PDF of those genes. The processing described herein is preferably performed in MATLAB™ (The Mathworks, Natick, Massachusetts) a programming language
5 and development environment for technical computing.

Iterative Processes

In practicing the invention, these processes can be iteratively repeated. For example, a nucleic acid (or, the nucleic acid) responsible for an altered or new xylanase phenotype is identified, re-isolated, again modified, re-tested for activity. This process can
10 be iteratively repeated until a desired phenotype is engineered. For example, an entire biochemical anabolic or catabolic pathway can be engineered into a cell, including, e.g., xylanase activity.

Similarly, if it is determined that a particular oligonucleotide has no affect at all on the desired trait (e.g., a new xylanase phenotype), it can be removed as a variable by synthesizing larger parental oligonucleotides that include the sequence to be removed. Since
15 incorporating the sequence within a larger sequence prevents any crossover events, there will no longer be any variation of this sequence in the progeny polynucleotides. This iterative practice of determining which oligonucleotides are most related to the desired trait, and which are unrelated, allows more efficient exploration all of the possible protein variants that
20 might be provide a particular trait or activity.

In vivo shuffling

In vivo shuffling of molecules is use in methods of the invention that provide variants of polypeptides of the invention, e.g., antibodies, xylanases, and the like. *In vivo* shuffling can be performed utilizing the natural property of cells to recombine multimers.
25 While recombination *in vivo* has provided the major natural route to molecular diversity, genetic recombination remains a relatively complex process that involves 1) the recognition of homologies; 2) strand cleavage, strand invasion, and metabolic steps leading to the production of recombinant chiasma; and finally 3) the resolution of chiasma into discrete recombined molecules. The formation of the chiasma requires the recognition of
30 homologous sequences.

In another aspect, the invention includes a method for producing a hybrid polynucleotide from at least a first polynucleotide and a second polynucleotide. The invention can be used to produce a hybrid polynucleotide by introducing at least a first

polynucleotide and a second polynucleotide which share at least one region of partial sequence homology (e.g., SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257 and combinations thereof) into a suitable host cell. The regions of partial sequence homology promote processes which result in sequence reorganization producing a hybrid polynucleotide. The term "hybrid polynucleotide", as used herein, is any nucleotide sequence which results from the method of the present invention and contains sequence from at least two original polynucleotide sequences. Such hybrid polynucleotides can result from intermolecular recombination events which promote sequence integration between DNA molecules. In addition, such hybrid polynucleotides can result from intramolecular reductive reassortment processes which utilize repeated sequences to alter a nucleotide sequence within a DNA molecule.

In *vivo* reassortment is focused on "inter-molecular" processes collectively referred to as "recombination" which in bacteria, is generally viewed as a "RecA-dependent" phenomenon. The invention can rely on recombination processes of a host cell to recombine and re-assort sequences, or the cells' ability to mediate reductive processes to decrease the complexity of quasi-repeated sequences in the cell by deletion. This process of "reductive reassortment" occurs by an "intra-molecular", RecA-independent process.

Therefore, in another aspect of the invention, novel polynucleotides can be generated by the process of reductive reassortment. The method involves the generation of constructs containing consecutive sequences (original encoding sequences), their insertion into an appropriate vector and their subsequent introduction into an appropriate host cell. The reassortment of the individual molecular identities occurs by combinatorial processes between the consecutive sequences in the construct possessing regions of homology, or between quasi-repeated units. The reassortment process recombines and/or reduces the complexity and extent of the repeated sequences and results in the production of novel molecular species. Various treatments may be applied to enhance the rate of reassortment. These could include treatment with ultra-violet light, or DNA damaging chemicals and/or the use of host cell lines displaying enhanced levels of "genetic instability". Thus the

reassortment process may involve homologous recombination or the natural property of quasi-repeated sequences to direct their own evolution.

Repeated or "quasi-repeated" sequences play a role in genetic instability. In the present invention, "quasi-repeats" are repeats that are not restricted to their original unit structure. Quasi-repeated units can be presented as an array of sequences in a construct; consecutive units of similar sequences. Once ligated, the junctions between the consecutive sequences become essentially invisible and the quasi-repetitive nature of the resulting construct is now continuous at the molecular level. The deletion process the cell performs to reduce the complexity of the resulting construct operates between the quasi-repeated sequences. The quasi-repeated units provide a practically limitless repertoire of templates upon which slippage events can occur. The constructs containing the quasi-repeats thus effectively provide sufficient molecular elasticity that deletion (and potentially insertion) events can occur virtually anywhere within the quasi-repetitive units.

When the quasi-repeated sequences are all ligated in the same orientation, for instance head to tail or vice versa, the cell cannot distinguish individual units. Consequently, the reductive process can occur throughout the sequences. In contrast, when for example, the units are presented head to head, rather than head to tail, the inversion delineates the endpoints of the adjacent unit so that deletion formation will favor the loss of discrete units. Thus, it is preferable with the present method that the sequences are in the same orientation. Random orientation of quasi-repeated sequences will result in the loss of reassortment efficiency, while consistent orientation of the sequences will offer the highest efficiency. However, while having fewer of the contiguous sequences in the same orientation decreases the efficiency, it may still provide sufficient elasticity for the effective recovery of novel molecules. Constructs can be made with the quasi-repeated sequences in the same orientation to allow higher efficiency.

Sequences can be assembled in a head to tail orientation using any of a variety of methods, including the following:

- a) Primers that include a poly-A head and poly-T tail which when made single-stranded would provide orientation can be utilized. This is accomplished by having the first few bases of the primers made from RNA and hence easily removed RNaseH.
- b) Primers that include unique restriction cleavage sites can be utilized. Multiple sites, a battery of unique sequences and repeated synthesis and ligation steps would be required.

- c) The inner few bases of the primer could be thiolated and an exonuclease used to produce properly tailed molecules.

The recovery of the re-assorted sequences relies on the identification of cloning vectors with a reduced repetitive index (RI). The re-assorted encoding sequences can then be recovered by amplification. The products are re-cloned and expressed. The recovery of cloning vectors with reduced RI can be affected by:

- 1) The use of vectors only stably maintained when the construct is reduced in complexity.
- 2) The physical recovery of shortened vectors by physical procedures. In this case, the cloning vector would be recovered using standard plasmid isolation procedures and size fractionated on either an agarose gel, or column with a low molecular weight cut off utilizing standard procedures.
- 3) The recovery of vectors containing interrupted genes which can be selected when insert size decreases.
- 4) The use of direct selection techniques with an expression vector and the appropriate selection.

Encoding sequences (for example, genes) from related organisms may demonstrate a high degree of homology and encode quite diverse protein products. These types of sequences are particularly useful in the present invention as quasi-repeats. However, while the examples illustrated below demonstrate the reassortment of nearly identical original encoding sequences (quasi-repeats), this process is not limited to such nearly identical repeats.

The following example demonstrates a method of the invention. Encoding nucleic acid sequences (quasi-repeats) derived from three (3) unique species are described. Each sequence encodes a protein with a distinct set of properties. Each of the sequences differs by a single or a few base pairs at a unique position in the sequence. The quasi-repeated sequences are separately or collectively amplified and ligated into random assemblies such that all possible permutations and combinations are available in the population of ligated molecules. The number of quasi-repeat units can be controlled by the assembly conditions. The average number of quasi-repeated units in a construct is defined as the repetitive index (RI).

Once formed, the constructs may, or may not be size fractionated on an agarose gel according to published protocols, inserted into a cloning vector and transfected

into an appropriate host cell. The cells are then propagated and “reductive reassortment” is effected. The rate of the reductive reassortment process may be stimulated by the introduction of DNA damage if desired. Whether the reduction in RI is mediated by deletion formation between repeated sequences by an “intra-molecular” mechanism, or mediated by recombination-like events through “inter-molecular” mechanisms is immaterial. The end result is a reassortment of the molecules into all possible combinations.

Optionally, the method comprises the additional step of screening the library members of the shuffled pool to identify individual shuffled library members having the ability to bind or otherwise interact, or catalyze a particular reaction (*e.g.*, such as catalytic domain of an enzyme) with a predetermined macromolecule, such as for example a proteinaceous receptor, an oligosaccharide, virion, or other predetermined compound or structure.

The polypeptides that are identified from such libraries can be used for therapeutic, diagnostic, research and related purposes (*e.g.*, catalysts, solutes for increasing osmolarity of an aqueous solution and the like) and/or can be subjected to one or more additional cycles of shuffling and/or selection.

In another aspect, it is envisioned that prior to or during recombination or reassortment, polynucleotides generated by the method of the invention can be subjected to agents or processes which promote the introduction of mutations into the original polynucleotides. The introduction of such mutations would increase the diversity of resulting hybrid polynucleotides and polypeptides encoded therefrom. The agents or processes which promote mutagenesis can include, but are not limited to: (+)-CC-1065, or a synthetic analog such as (+)-CC-1065-(N3-Adenine (*See* Sun and Hurley, (1992); an N-acetylated or deacetylated 4'-fluoro-4-aminobiphenyl adduct capable of inhibiting DNA synthesis (*See*, for example, van de Poll *et al.* (1992)); or a N-acetylated or deacetylated 4-aminobiphenyl adduct capable of inhibiting DNA synthesis (*See* also, van de Poll *et al.* (1992), pp. 751-758); trivalent chromium, a trivalent chromium salt, a polycyclic aromatic hydrocarbon (PAH) DNA adduct capable of inhibiting DNA replication, such as 7-bromomethyl-benz[*a*]anthracene (“BMA”), tris(2,3-dibromopropyl)phosphate (“Tris-BP”), 1,2-dibromo-3-chloropropane (“DBCP”), 2-bromoacrolein (2BA), benzo[*a*]pyrene-7,8-dihydrodiol-9-10-epoxide (“BPDE”), a platinum(II) halogen salt, N-hydroxy-2-amino-3-methylimidazo[4,5-*f*]quinoline (“N-hydroxy-IQ”) and N-hydroxy-2-amino-1-methyl-6-phenylimidazo[4,5-*f*]pyridine (“N-hydroxy-PhIP”). Exemplary means for slowing or halting PCR amplification consist of UV light (+)-CC-1065 and (+)-CC-1065-(N3-Adenine). Particularly encompassed

means are DNA adducts or polynucleotides comprising the DNA adducts from the polynucleotides or polynucleotides pool, which can be released or removed by a process including heating the solution comprising the polynucleotides prior to further processing.

In another aspect the invention is directed to a method of producing
5 recombinant proteins having biological activity by treating a sample comprising double-stranded template polynucleotides encoding a wild-type protein under conditions according to the invention which provide for the production of hybrid or re-assorted polynucleotides.

Producing sequence variants

The invention also provides additional methods for making sequence variants
10 of the nucleic acid (e.g., xylanase) sequences of the invention. The invention also provides additional methods for isolating xylanases using the nucleic acids and polypeptides of the invention. In one aspect, the invention provides for variants of a xylanase coding sequence (e.g., a gene, cDNA or message) of the invention, which can be altered by any means, including, e.g., random or stochastic methods, or, non-stochastic, or "directed evolution,"
15 methods, as described above.

The isolated variants may be naturally occurring. Variant can also be created *in vitro*. Variants may be created using genetic engineering techniques such as site directed mutagenesis, random chemical mutagenesis, Exonuclease III deletion procedures, and standard cloning techniques. Alternatively, such variants, fragments, analogs, or derivatives
20 may be created using chemical synthesis or modification procedures. Other methods of making variants are also familiar to those skilled in the art. These include procedures in which nucleic acid sequences obtained from natural isolates are modified to generate nucleic acids which encode polypeptides having characteristics which enhance their value in industrial or laboratory applications. In such procedures, a large number of variant sequences
25 having one or more nucleotide differences with respect to the sequence obtained from the natural isolate are generated and characterized. These nucleotide differences can result in amino acid changes with respect to the polypeptides encoded by the nucleic acids from the natural isolates.

For example, variants may be created using error prone PCR. In error prone
30 PCR, PCR is performed under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. Error prone PCR is described, e.g., in Leung, D.W., et al., Technique, 1:11-15, 1989) and Caldwell, R. C. & Joyce G.F., PCR Methods Applic., 2:28-33, 1992. Briefly, in

such procedures, nucleic acids to be mutagenized are mixed with PCR primers, reaction buffer, MgCl₂, MnCl₂, Taq polymerase and an appropriate concentration of dNTPs for achieving a high rate of point mutation along the entire length of the PCR product. For example, the reaction may be performed using 20 fmoles of nucleic acid to be mutagenized, 5 30 pmole of each PCR primer, a reaction buffer comprising 50mM KCl, 10mM Tris HCl (pH 8.3) and 0.01% gelatin, 7mM MgCl₂, 0.5mM MnCl₂, 5 units of Taq polymerase, 0.2mM dGTP, 0.2mM dATP, 1mM dCTP, and 1mM dTTP. PCR may be performed for 30 cycles of 94°C for 1 min, 45°C for 1 min, and 72°C for 1 min. However, it will be appreciated that these parameters may be varied as appropriate. The mutagenized nucleic acids are cloned 10 into an appropriate vector and the activities of the polypeptides encoded by the mutagenized nucleic acids are evaluated.

Variants may also be created using oligonucleotide directed mutagenesis to generate site-specific mutations in any cloned DNA of interest. Oligonucleotide mutagenesis is described, e.g., in Reidhaar-Olson (1988) Science 241:53-57. Briefly, in such procedures a 15 plurality of double stranded oligonucleotides bearing one or more mutations to be introduced into the cloned DNA are synthesized and inserted into the cloned DNA to be mutagenized. Clones containing the mutagenized DNA are recovered and the activities of the polypeptides they encode are assessed.

Another method for generating variants is assembly PCR. Assembly PCR 20 involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions occur in parallel in the same vial, with the products of one reaction priming the products of another reaction. Assembly PCR is described in, e.g., U.S. Patent No. 5,965,408.

Still another method of generating variants is sexual PCR mutagenesis. In 25 sexual PCR mutagenesis, forced homologous recombination occurs between DNA molecules of different but highly related DNA sequence *in vitro*, as a result of random fragmentation of the DNA molecule based on sequence homology, followed by fixation of the crossover by primer extension in a PCR reaction. Sexual PCR mutagenesis is described, e.g., in Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751. Briefly, in such procedures a plurality 30 of nucleic acids to be recombined are digested with DNase to generate fragments having an average size of 50-200 nucleotides. Fragments of the desired average size are purified and resuspended in a PCR mixture. PCR is conducted under conditions which facilitate recombination between the nucleic acid fragments. For example, PCR may be performed by resuspending the purified fragments at a concentration of 10-30ng/μl in a solution of 0.2mM

of each dNTP, 2.2mM MgCl₂, 50mM KCL, 10mM Tris HCl, pH 9.0, and 0.1% Triton X-100. 2.5 units of Taq polymerase per 100:1 of reaction mixture is added and PCR is performed using the following regime: 94°C for 60 seconds, 94°C for 30 seconds, 50-55°C for 30 seconds, 72°C for 30 seconds (30-45 times) and 72°C for 5 minutes. However, it will be appreciated that these parameters may be varied as appropriate. In some aspects, oligonucleotides may be included in the PCR reactions. In other aspects, the Klenow fragment of DNA polymerase I may be used in a first set of PCR reactions and Taq polymerase may be used in a subsequent set of PCR reactions. Recombinant sequences are isolated and the activities of the polypeptides they encode are assessed.

10 Variants may also be created by *in vivo* mutagenesis. In some aspects, random mutations in a sequence of interest are generated by propagating the sequence of interest in a bacterial strain, such as an E. coli strain, which carries mutations in one or more of the DNA repair pathways. Such "mutator" strains have a higher random mutation rate than that of a wild-type parent. Propagating the DNA in one of these strains will eventually generate random mutations within the DNA. Mutator strains suitable for use for *in vivo* mutagenesis are described in PCT Publication No. WO 91/16427, published October 31, 1991, entitled "Methods for Phenotype Creation from Multiple Gene Populations".

15 Variants may also be generated using cassette mutagenesis. In cassette mutagenesis a small region of a double stranded DNA molecule is replaced with a synthetic oligonucleotide "cassette" that differs from the native sequence. The oligonucleotide often contains completely and/or partially randomized native sequence.

Recursive ensemble mutagenesis may also be used to generate variants. Recursive ensemble mutagenesis is an algorithm for protein engineering (protein mutagenesis) developed to produce diverse populations of phenotypically related mutants whose members differ in amino acid sequence. This method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. Recursive ensemble mutagenesis is described in Arkin, A.P. and Youvan, D.C., PNAS, USA, 89:7811-7815, 1992.

20 In some aspects, variants are created using exponential ensemble mutagenesis. Exponential ensemble mutagenesis is a process for generating combinatorial libraries with a high percentage of unique and functional mutants, wherein small groups of residues are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. Exponential ensemble mutagenesis is described in Delegrave, S. and Youvan, D.C., Biotechnology Research, 11:1548-1552, 1993. Random and site-directed

mutagenesis are described in Arnold, F.H., Current Opinion in Biotechnology, 4:450-455, 1993.

In some aspects, the variants are created using shuffling procedures wherein portions of a plurality of nucleic acids which encode distinct polypeptides are fused together to create chimeric nucleic acid sequences which encode chimeric polypeptides as described in
5 U.S. Patent No. 5,965,408, filed July 9, 1996, entitled, "Method of DNA Reassembly by Interrupting Synthesis" and U.S. Patent No. 5,939,250, filed May 22, 1996, entitled, "Production of Enzymes Having Desired Activities by Mutagenesis.

The variants of the polypeptides of Group B amino acid sequences may be
10 variants in which one or more of the amino acid residues of the polypeptides of the Group B amino acid sequences are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code.

Conservative substitutions are those that substitute a given amino acid in a
15 polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the following replacements: replacements of an aliphatic amino acid such as Alanine, Valine, Leucine and Isoleucine with another aliphatic amino acid; replacement of a Serine with a Threonine or vice versa; replacement of an acidic residue such as Aspartic acid and Glutamic acid with another acidic residue; replacement of a residue bearing an amide
20 group, such as Asparagine and Glutamine, with another residue bearing an amide group; exchange of a basic residue such as Lysine and Arginine with another basic residue; and replacement of an aromatic residue such as Phenylalanine, Tyrosine with another aromatic residue.

Other variants are those in which one or more of the amino acid residues of the
25 polypeptides of the Group B amino acid sequences includes a substituent group.

Still other variants are those in which the polypeptide is associated with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol).

Additional variants are those in which additional amino acids are fused to the
30 polypeptide, such as a leader sequence, a secretory sequence, a proprotein sequence or a sequence which facilitates purification, enrichment, or stabilization of the polypeptide.

In some aspects, the fragments, derivatives and analogs retain the same biological function or activity as the polypeptides of Group B amino acid sequences and sequences substantially identical thereto. In other aspects, the fragment, derivative, or analog

includes a proprotein, such that the fragment, derivative, or analog can be activated by cleavage of the proprotein portion to produce an active polypeptide.

Optimizing codons to achieve high levels of protein expression in host cells

The invention provides methods for modifying xylanase-encoding nucleic acids to modify codon usage. In one aspect, the invention provides methods for modifying codons in a nucleic acid encoding a xylanase to increase or decrease its expression in a host cell. The invention also provides nucleic acids encoding a xylanase modified to increase its expression in a host cell, xylanase so modified, and methods of making the modified xylanases. The method comprises identifying a “non-preferred” or a “less preferred” codon in xylanase-encoding nucleic acid and replacing one or more of these non-preferred or less preferred codons with a “preferred codon” encoding the same amino acid as the replaced codon and at least one non-preferred or less preferred codon in the nucleic acid has been replaced by a preferred codon encoding the same amino acid. A preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell.

Host cells for expressing the nucleic acids, expression cassettes and vectors of the invention include bacteria, yeast, fungi, plant cells, insect cells and mammalian cells. Thus, the invention provides methods for optimizing codon usage in all of these cells, codon-altered nucleic acids and polypeptides made by the codon-altered nucleic acids. Exemplary host cells include gram negative bacteria, such as *Escherichia coli* and *Pseudomonas fluorescens*; gram positive bacteria, such as *Streptomyces diversa*, *Lactobacillus gasseri*, *Lactococcus lactis*, *Lactococcus cremoris*, *Bacillus subtilis*. Exemplary host cells also include eukaryotic organisms, e.g., various yeast, such as *Saccharomyces* sp., including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, and *Kluyveromyces lactis*, *Hansenula polymorpha*, *Aspergillus niger*, and mammalian cells and cell lines and insect cells and cell lines. Thus, the invention also includes nucleic acids and polypeptides optimized for expression in these organisms and species.

For example, the codons of a nucleic acid encoding a xylanase isolated from a bacterial cell are modified such that the nucleic acid is optimally expressed in a bacterial cell different from the bacteria from which the xylanase was derived, a yeast, a fungi, a plant cell, an insect cell or a mammalian cell. Methods for optimizing codons are well known in the art, see, e.g., U.S. Patent No. 5,795,737; Baca (2000) Int. J. Parasitol. 30:113-118; Hale (1998) Protein Expr. Purif. 12:185-188; Narum (2001) Infect. Immun. 69:7250-7253. See also

Narum (2001) *Infect. Immun.* 69:7250-7253, describing optimizing codons in mouse systems; Outchkourov (2002) *Protein Expr. Purif.* 24:18-24, describing optimizing codons in yeast; Feng (2000) *Biochemistry* 39:15399-15409, describing optimizing codons in *E. coli*; Humphreys (2000) *Protein Expr. Purif.* 20:252-264, describing optimizing codon usage that affects secretion in *E. coli*.

Transgenic non-human animals

The invention provides transgenic non-human animals comprising a nucleic acid, a polypeptide (e.g., a xylanase), an expression cassette or vector or a transfected or transformed cell of the invention. The invention also provides methods of making and using these transgenic non-human animals.

The transgenic non-human animals can be, e.g., goats, rabbits, sheep, pigs, cows, rats and mice, comprising the nucleic acids of the invention. These animals can be used, e.g., as *in vivo* models to study xylanase activity, or, as models to screen for agents that change the xylanase activity *in vivo*. The coding sequences for the polypeptides to be expressed in the transgenic non-human animals can be designed to be constitutive, or, under the control of tissue-specific, developmental-specific or inducible transcriptional regulatory factors. Transgenic non-human animals can be designed and generated using any method known in the art; see, e.g., U.S. Patent Nos. 6,211,428; 6,187,992; 6,156,952; 6,118,044; 6,111,166; 6,107,541; 5,959,171; 5,922,854; 5,892,070; 5,880,327; 5,891,698; 5,639,940; 5,573,933; 5,387,742; 5,087,571, describing making and using transformed cells and eggs and transgenic mice, rats, rabbits, sheep, pigs and cows. See also, e.g., Pollock (1999) *J. Immunol. Methods* 231:147-157, describing the production of recombinant proteins in the milk of transgenic dairy animals; Baguisi (1999) *Nat. Biotechnol.* 17:456-461, demonstrating the production of transgenic goats. U.S. Patent No. 6,211,428, describes making and using transgenic non-human mammals which express in their brains a nucleic acid construct comprising a DNA sequence. U.S. Patent No. 5,387,742, describes injecting cloned recombinant or synthetic DNA sequences into fertilized mouse eggs, implanting the injected eggs in pseudo-pregnant females, and growing to term transgenic mice whose cells express proteins related to the pathology of Alzheimer's disease. U.S. Patent No. 6,187,992, describes making and using a transgenic mouse whose genome comprises a disruption of the gene encoding amyloid precursor protein (APP).

"Knockout animals" can also be used to practice the methods of the invention. For example, in one aspect, the transgenic or modified animals of the invention comprise a "knockout animal," e.g., a "knockout mouse," engineered not to express an endogenous gene,

which is replaced with a gene expressing a xylanase of the invention, or, a fusion protein comprising a xylanase of the invention.

Transgenic Plants and Seeds

The invention provides transgenic plants and seeds comprising a nucleic acid,
5 a polypeptide (e.g., a xylanase), an expression cassette or vector or a transfected or transformed cell of the invention. The invention also provides plant products, e.g., oils, seeds, leaves, extracts and the like, comprising a nucleic acid and/or a polypeptide (e.g., a xylanase) of the invention. The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot). The invention also provides methods of making and using
10 these transgenic plants and seeds. The transgenic plant or plant cell expressing a polypeptide of the present invention may be constructed in accordance with any method known in the art. See, for example, U.S. Patent No. 6,309,872.

Nucleic acids and expression constructs of the invention can be introduced into a plant cell by any means. For example, nucleic acids or expression constructs can be
15 introduced into the genome of a desired plant host, or, the nucleic acids or expression constructs can be episomes. Introduction into the genome of a desired plant can be such that the host's xylanase production is regulated by endogenous transcriptional or translational control elements. The invention also provides "knockout plants" where insertion of gene sequence by, e.g., homologous recombination, has disrupted the expression of the
20 endogenous gene. Means to generate "knockout" plants are well-known in the art, see, e.g., Strepp (1998) Proc Natl. Acad. Sci. USA 95:4368-4373; Miao (1995) Plant J 7:359-365. See discussion on transgenic plants, below.

The nucleic acids of the invention can be used to confer desired traits on essentially any plant, e.g., on starch-producing plants, such as potato, wheat, rice, barley, and
25 the like. Nucleic acids of the invention can be used to manipulate metabolic pathways of a plant in order to optimize or alter host's expression of xylanase. The can change xylanase activity in a plant. Alternatively, a xylanase of the invention can be used in production of a transgenic plant to produce a compound not naturally produced by that plant. This can lower production costs or create a novel product.

30 In one aspect, the first step in production of a transgenic plant involves making an expression construct for expression in a plant cell. These techniques are well known in the art. They can include selecting and cloning a promoter, a coding sequence for facilitating efficient binding of ribosomes to mRNA and selecting the appropriate gene terminator sequences. One exemplary constitutive promoter is CaMV35S, from the cauliflower mosaic

virus, which generally results in a high degree of expression in plants. Other promoters are more specific and respond to cues in the plant's internal or external environment. An exemplary light-inducible promoter is the promoter from the *cab* gene, encoding the major chlorophyll *a/b* binding protein.

5 In one aspect, the nucleic acid is modified to achieve greater expression in a plant cell. For example, a sequence of the invention is likely to have a higher percentage of A-T nucleotide pairs compared to that seen in a plant, some of which prefer G-C nucleotide pairs. Therefore, A-T nucleotides in the coding sequence can be substituted with G-C nucleotides without significantly changing the amino acid sequence to enhance production of
10 the gene product in plant cells.

 Selectable marker gene can be added to the gene construct in order to identify plant cells or tissues that have successfully integrated the transgene. This may be necessary because achieving incorporation and expression of genes in plant cells is a rare event, occurring in just a few percent of the targeted tissues or cells. Selectable marker genes
15 encode proteins that provide resistance to agents that are normally toxic to plants, such as antibiotics or herbicides. Only plant cells that have integrated the selectable marker gene will survive when grown on a medium containing the appropriate antibiotic or herbicide. As for other inserted genes, marker genes also require promoter and termination sequences for proper function.

20 In one aspect, making transgenic plants or seeds comprises incorporating sequences of the invention and, optionally, marker genes into a target expression construct (e.g., a plasmid), along with positioning of the promoter and the terminator sequences. This can involve transferring the modified gene into the plant through a suitable method. For example, a construct may be introduced directly into the genomic DNA of the plant cell using
25 techniques such as electroporation and microinjection of plant cell protoplasts, or the constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. For example, see, e.g., Christou (1997) *Plant Mol. Biol.* 35:197-203; Pawlowski (1996) *Mol. Biotechnol.* 6:17-30; Klein (1987) *Nature* 327:70-73; Takumi (1997) *Genes Genet. Syst.* 72:63-69, discussing use of particle bombardment to introduce transgenes
30 into wheat; and Adam (1997) *supra*, for use of particle bombardment to introduce YACs into plant cells. For example, Rinehart (1997) *supra*, used particle bombardment to generate transgenic cotton plants. Apparatus for accelerating particles is described U.S. Pat. No. 5,015,580; and, the commercially available BioRad (Biolistics) PDS-2000 particle

acceleration instrument; see also, John, U.S. Patent No. 5,608,148; and Ellis, U.S. Patent No. 5, 681,730, describing particle-mediated transformation of gymnosperms.

In one aspect, protoplasts can be immobilized and injected with a nucleic acids, e.g., an expression construct. Although plant regeneration from protoplasts is not easy with cereals, plant regeneration is possible in legumes using somatic embryogenesis from protoplast derived callus. Organized tissues can be transformed with naked DNA using gene gun technique, where DNA is coated on tungsten microprojectiles, shot 1/100th the size of cells, which carry the DNA deep into cells and organelles. Transformed tissue is then induced to regenerate, usually by somatic embryogenesis. This technique has been successful in several cereal species including maize and rice.

Nucleic acids, e.g., expression constructs, can also be introduced in to plant cells using recombinant viruses. Plant cells can be transformed using viral vectors, such as, e.g., tobacco mosaic virus derived vectors (Rouwendal (1997) *Plant Mol. Biol.* 33:989-999), see Porta (1996) "Use of viral replicons for the expression of genes in plants," *Mol. Biotechnol.* 5:209-221.

Alternatively, nucleic acids, e.g., an expression construct, can be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. *Agrobacterium tumefaciens*-mediated transformation techniques, including disarming and use of binary vectors, are well described in the scientific literature. See, e.g., Horsch (1984) *Science* 233:496-498; Fraley (1983) *Proc. Natl. Acad. Sci. USA* 80:4803 (1983); *Gene Transfer to Plants*, Potrykus, ed. (Springer-Verlag, Berlin 1995). The DNA in an *A. tumefaciens* cell is contained in the bacterial chromosome as well as in another structure known as a Ti (tumor-inducing) plasmid. The Ti plasmid contains a stretch of DNA termed T-DNA (~20 kb long) that is transferred to the plant cell in the infection process and a series of vir (virulence) genes that direct the infection process. *A. tumefaciens* can only infect a plant through wounds: when a plant root or stem is wounded it gives off certain chemical signals, in response to which, the vir genes of *A. tumefaciens* become activated and direct a series of events necessary for the transfer of the T-DNA from the Ti plasmid to the plant's chromosome. The T-DNA then enters the plant cell through the wound. One speculation is that the T-DNA waits until the plant DNA is being replicated or transcribed, then inserts itself into the exposed plant DNA. In order to use *A. tumefaciens* as a transgene vector, the tumor-inducing section of T-DNA have to be removed, while retaining the T-DNA border regions

and the vir genes. The transgene is then inserted between the T-DNA border regions, where it is transferred to the plant cell and becomes integrated into the plant's chromosomes.

The invention provides for the transformation of monocotyledonous plants using the nucleic acids of the invention, including important cereals, see Hiei (1997) Plant Mol. Biol. 35:205-218. See also, e.g., Horsch, Science (1984) 233:496; Fraley (1983) Proc. Natl. Acad. Sci USA 80:4803; Thykjaer (1997) supra; Park (1996) Plant Mol. Biol. 32:1135-1148, discussing T-DNA integration into genomic DNA. See also D'Halluin, U.S. Patent No. 5,712,135, describing a process for the stable integration of a DNA comprising a gene that is functional in a cell of a cereal, or other monocotyledonous plant.

In one aspect, the third step can involve selection and regeneration of whole plants capable of transmitting the incorporated target gene to the next generation. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker that has been introduced together with the desired nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans et al., *Protoplasts Isolation and Culture, Handbook of Plant Cell Culture*, pp. 124-176, MacMillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee (1987) Ann. Rev. of Plant Phys. 38:467-486. To obtain whole plants from transgenic tissues such as immature embryos, they can be grown under controlled environmental conditions in a series of media containing nutrients and hormones, a process known as tissue culture. Once whole plants are generated and produce seed, evaluation of the progeny begins.

After the expression cassette is stably incorporated in transgenic plants, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed. Since transgenic expression of the nucleic acids of the invention leads to phenotypic changes, plants comprising the recombinant nucleic acids of the invention can be sexually crossed with a second plant to obtain a final product. Thus, the seed of the invention can be derived from a cross between two transgenic plants of the invention, or a cross between a plant of the invention and another plant. The desired effects (e.g., expression of the polypeptides of the invention to produce a plant in which flowering behavior is altered) can be enhanced when both parental plants express the polypeptides (e.g., a xylanase) of the invention. The desired effects can be passed to future plant generations by standard propagation means.

The nucleic acids and polypeptides of the invention are expressed in or inserted in any plant or seed. Transgenic plants of the invention can be dicotyledonous or monocotyledonous. Examples of monocot transgenic plants of the invention are grasses, such as meadow grass (blue grass, *Poa*), forage grass such as festuca, lolium, temperate grass, such as *Agrostis*, and cereals, e.g., wheat, oats, rye, barley, rice, sorghum, and maize (corn). Examples of dicot transgenic plants of the invention are tobacco, legumes, such as lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family *Brassicaceae*), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*. Thus, the transgenic plants and seeds of the invention include a broad range of plants, including, but not limited to, species from the genera *Anacardium*, *Arachis*, *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Carthamus*, *Cocos*, *Coffea*, *Cucumis*, *Cucurbita*, *Daucus*, *Elaeis*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Olea*, *Oryza*, *Panicum*, *Pennisetum*, *Persea*, *Phaseolus*, *Pistachia*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Ricinus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Theobromus*, *Trigonella*, *Triticum*, *Vicia*, *Vitis*, *Vigna*, and *Zea*.

In alternative embodiments, the nucleic acids of the invention are expressed in plants which contain fiber cells, including, e.g., cotton, silk cotton tree (Kapok, *Ceiba pentandra*), desert willow, creosote bush, winterfat, balsa, ramie, kenaf, hemp, roselle, jute, sisal abaca and flax. In alternative embodiments, the transgenic plants of the invention can be members of the genus *Gossypium*, including members of any *Gossypium* species, such as *G. arboreum*, *G. herbaceum*, *G. barbadense*, and *G. hirsutum*.

The invention also provides for transgenic plants to be used for producing large amounts of the polypeptides (e.g., a xylanase or antibody) of the invention. For example, see Palmgren (1997) Trends Genet. 13:348; Chong (1997) Transgenic Res. 6:289-296 (producing human milk protein beta-casein in transgenic potato plants using an auxin-inducible, bidirectional mannopine synthase (*mas1',2'*) promoter with *Agrobacterium tumefaciens*-mediated leaf disc transformation methods).

Using known procedures, one of skill can screen for plants of the invention by detecting the increase or decrease of transgene mRNA or protein in transgenic plants. Means for detecting and quantitation of mRNAs or proteins are well known in the art.

Polypeptides and peptides

In one aspect, the invention provides isolated or recombinant polypeptides having a sequence identity (e.g., at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or complete (100%) sequence identity) to an exemplary sequence of the invention, e.g., proteins having a sequence as set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198, SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228, SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:244, SEQ ID NO:246, SEQ ID NO:248, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256, SEQ ID NO:258, SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:268, SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276, SEQ ID NO:278,

SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286, SEQ ID NO:288,
SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296, SEQ ID NO:298,
SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306, SEQ ID NO:308,
SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316, SEQ ID NO:318,
5 SEQ ID NO:320, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328,
SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:338,
SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348,
SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358,
SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368,
10 SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378 or
SEQ ID NO:380. In one aspect, the polypeptide has a xylanase activity, e.g., can hydrolyze a
glycosidic bond in a polysaccharide, e.g., a xylan. In one aspect, the polypeptide has a
xylanase activity comprising catalyzing hydrolysis of internal β -1,4-xylosidic linkages. In
one aspect, the xylanase activity comprises an endo-1,4-beta-xylanase activity. In one aspect,
15 the xylanase activity comprises hydrolyzing a xylan to produce a smaller molecular weight
xylose and xylo-oligomer. In one aspect, the xylan comprises an arabinoxylan, such as a
water soluble arabinoxylan.

The polypeptides of the invention include xylanases in an active or inactive
form. For example, the polypeptides of the invention include proproteins before
20 "maturation" or processing of prepro sequences, e.g., by a proprotein-processing enzyme,
such as a proprotein convertase to generate an "active" mature protein. The polypeptides of
the invention include xylanases inactive for other reasons, e.g., before "activation" by a post-
translational processing event, e.g., an endo- or exo-peptidase or proteinase action, a
phosphorylation event, an amidation, a glycosylation or a sulfation, a dimerization event, and
25 the like. The polypeptides of the invention include all active forms, including active
subsequences, e.g., catalytic domains or active sites, of the xylanase.

Methods for identifying "prepro" domain sequences and signal sequences are
well known in the art, see, e.g., Van de Ven (1993) Crit. Rev. Oncog. 4(2):115-136. For
example, to identify a prepro sequence, the protein is purified from the extracellular space
30 and the N-terminal protein sequence is determined and compared to the unprocessed form.

The invention includes polypeptides with or without a signal sequence and/or
a prepro sequence. The invention includes polypeptides with heterologous signal sequences
and/or prepro sequences. The prepro sequence (including a sequence of the invention used as
a heterologous prepro domain) can be located on the amino terminal or the carboxy terminal

end of the protein. The invention also includes isolated or recombinant signal sequences, prepro sequences and catalytic domains (e.g., "active sites") comprising sequences of the invention.

The percent sequence identity can be over the full length of the polypeptide, or, the identity can be over a region of at least about 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 or more residues. Polypeptides of the invention can also be shorter than the full length of exemplary polypeptides. In alternative aspects, the invention provides polypeptides (peptides, fragments) ranging in size between about 5 and the full length of a polypeptide, e.g., an enzyme, such as a xylanase; exemplary sizes being of about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, or more residues, e.g., contiguous residues of an exemplary xylanase of the invention.

Peptides of the invention (e.g., a subsequence of an exemplary polypeptide of the invention) can be useful as, e.g., labeling probes, antigens, toleragens, motifs, xylanase active sites (e.g., "catalytic domains"), signal sequences and/or prepro domains.

Polypeptides and peptides of the invention can be isolated from natural sources, be synthetic, or be recombinantly generated polypeptides. Peptides and proteins can be recombinantly expressed *in vitro* or *in vivo*. The peptides and polypeptides of the invention can be made and isolated using any method known in the art. Polypeptide and peptides of the invention can also be synthesized, whole or in part, using chemical methods well known in the art. See e.g., Caruthers (1980) Nucleic Acids Res. Symp. Ser. 215-223; Horn (1980) Nucleic Acids Res. Symp. Ser. 225-232; Banga, A.K., Therapeutic Peptides and Proteins, Formulation, Processing and Delivery Systems (1995) Technomic Publishing Co., Lancaster, PA. For example, peptide synthesis can be performed using various solid-phase techniques (see e.g., Roberge (1995) Science 269:202; Merrifield (1997) Methods Enzymol. 289:3-13) and automated synthesis may be achieved, e.g., using the ABI 431A Peptide Synthesizer (Perkin Elmer) in accordance with the instructions provided by the manufacturer.

The peptides and polypeptides of the invention can also be glycosylated. The glycosylation can be added post-translationally either chemically or by cellular biosynthetic mechanisms, wherein the later incorporates the use of known glycosylation motifs, which can be native to the sequence or can be added as a peptide or added in the nucleic acid coding sequence. The glycosylation can be O-linked or N-linked.

The peptides and polypeptides of the invention, as defined above, include all "mimetic" and "peptidomimetic" forms. The terms "mimetic" and "peptidomimetic" refer to

a synthetic chemical compound which has substantially the same structural and/or functional characteristics of the polypeptides of the invention. The mimetic can be either entirely composed of synthetic, non-natural analogues of amino acids, or, is a chimeric molecule of partly natural peptide amino acids and partly non-natural analogs of amino acids. The
5 mimetic can also incorporate any amount of natural amino acid conservative substitutions as long as such substitutions also do not substantially alter the mimetic's structure and/or activity. As with polypeptides of the invention which are conservative variants, routine experimentation will determine whether a mimetic is within the scope of the invention, i.e., that its structure and/or function is not substantially altered. Thus, in one aspect, a mimetic
10 composition is within the scope of the invention if it has a xylanase activity.

Polypeptide mimetic compositions of the invention can contain any combination of non-natural structural components. In alternative aspect, mimetic compositions of the invention include one or all of the following three structural groups: a) residue linkage groups other than the natural amide bond ("peptide bond") linkages; b) non-
15 natural residues in place of naturally occurring amino acid residues; or c) residues which induce secondary structural mimicry, i.e., to induce or stabilize a secondary structure, e.g., a beta turn, gamma turn, beta sheet, alpha helix conformation, and the like. For example, a polypeptide of the invention can be characterized as a mimetic when all or some of its residues are joined by chemical means other than natural peptide bonds. Individual
20 peptidomimetic residues can be joined by peptide bonds, other chemical bonds or coupling means, such as, e.g., glutaraldehyde, N-hydroxysuccinimide esters, bifunctional maleimides, N,N'-dicyclohexylcarbodiimide (DCC) or N,N'-diisopropylcarbodiimide (DIC). Linking groups that can be an alternative to the traditional amide bond ("peptide bond") linkages include, e.g., ketomethylene (e.g., -C(=O)-CH₂- for -C(=O)-NH-), aminomethylene (CH₂-
25 NH), ethylene, olefin (CH=CH), ether (CH₂-O), thioether (CH₂-S), tetrazole (CN₄-), thiazole, retroamide, thioamide, or ester (see, e.g., Spatola (1983) in Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, "Peptide Backbone Modifications," Marcell Dekker, NY).

A polypeptide of the invention can also be characterized as a mimetic by
30 containing all or some non-natural residues in place of naturally occurring amino acid residues. Non-natural residues are well described in the scientific and patent literature; a few exemplary non-natural compositions useful as mimetics of natural amino acid residues and guidelines are described below. Mimetics of aromatic amino acids can be generated by replacing by, e.g., D- or L- naphylalanine; D- or L- phenylglycine; D- or L-2 thieneylalanine;

D- or L-1, -2, 3-, or 4- pyrenylalanine; D- or L-3 thienylalanine; D- or L-(2-pyridinyl)-alanine; D- or L-(3-pyridinyl)-alanine; D- or L-(2-pyrazinyl)-alanine; D- or L-(4-isopropyl)-phenylglycine; D-(trifluoromethyl)-phenylglycine; D-(trifluoromethyl)-phenylalanine; D-p-fluoro-phenylalanine; D- or L-p-biphenylphenylalanine; D- or L-p-methoxy-

5 biphenylphenylalanine; D- or L-2-indole(alkyl)alanines; and, D- or L-alkylainines, where alkyl can be substituted or unsubstituted methyl, ethyl, propyl, hexyl, butyl, pentyl, isopropyl, iso-butyl, sec-isotyl, iso-pentyl, or a non-acidic amino acids. Aromatic rings of a non-natural amino acid include, e.g., thiazolyl, thiophenyl, pyrazolyl, benzimidazolyl, naphthyl, furanyl, pyrrolyl, and pyridyl aromatic rings.

10 Mimetics of acidic amino acids can be generated by substitution by, e.g., non-carboxylate amino acids while maintaining a negative charge; (phosphono)alanine; sulfated threonine. Carboxyl side groups (e.g., aspartyl or glutamyl) can also be selectively modified by reaction with carbodiimides (R'-N-C-N-R') such as, e.g., 1-cyclohexyl-3(2-morpholinyl-(4-ethyl) carbodiimide or 1-ethyl-3(4-azonia- 4,4- dimetholpentyl) carbodiimide. Aspartyl or
15 glutamyl can also be converted to asparaginyl and glutaminyl residues by reaction with ammonium ions. Mimetics of basic amino acids can be generated by substitution with, e.g., (in addition to lysine and arginine) the amino acids ornithine, citrulline, or (guanidino)-acetic acid, or (guanidino)alkyl-acetic acid, where alkyl is defined above. Nitrile derivative (e.g., containing the CN-moiety in place of COOH) can be substituted for asparagine or glutamine.
20 Asparaginyl and glutaminyl residues can be deaminated to the corresponding aspartyl or glutamyl residues. Arginine residue mimetics can be generated by reacting arginyl with, e.g., one or more conventional reagents, including, e.g., phenylglyoxal, 2,3-butanedione, 1,2-cyclo-hexanedione, or ninhydrin, preferably under alkaline conditions. Tyrosine residue mimetics can be generated by reacting tyrosyl with, e.g., aromatic diazonium compounds or
25 tetranitromethane. N-acetylimidizol and tetranitromethane can be used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively. Cysteine residue mimetics can be generated by reacting cysteinyl residues with, e.g., alpha-haloacetates such as 2-chloroacetic acid or chloroacetamide and corresponding amines; to give carboxymethyl or carboxyamidomethyl derivatives. Cysteine residue mimetics can also be generated by
30 reacting cysteinyl residues with, e.g., bromo-trifluoroacetone, alpha-bromo-beta-(5-imidozoyl) propionic acid; chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide; methyl 2-pyridyl disulfide; p-chloromercuribenzoate; 2-chloromercuri-4-nitrophenol; or, chloro-7-nitrobenzo-oxa-1,3-diazole. Lysine mimetics can be generated (and amino terminal residues can be altered) by reacting lysinyl with, e.g., succinic or other

carboxylic acid anhydrides. Lysine and other alpha-amino-containing residue mimetics can also be generated by reaction with imidoesters, such as methyl picolinimide, pyridoxal phosphate, pyridoxal, chloroborohydride, trinitro-benzenesulfonic acid, O-methylisourea, 2,4-pentanedione, and transamidase-catalyzed reactions with glyoxylate. Mimetics of methionine
5 can be generated by reaction with, e.g., methionine sulfoxide. Mimetics of proline include, e.g., pipercolic acid, thiazolidine carboxylic acid, 3- or 4- hydroxy proline, dehydropoline, 3- or 4-methylproline, or 3,3,-dimethylproline. Histidine residue mimetics can be generated by reacting histidyl with, e.g., diethylprocarbonate or para-bromophenacyl bromide. Other mimetics include, e.g., those generated by hydroxylation of proline and lysine;
10 phosphorylation of the hydroxyl groups of seryl or threonyl residues; methylation of the alpha-amino groups of lysine, arginine and histidine; acetylation of the N-terminal amine; methylation of main chain amide residues or substitution with N-methyl amino acids; or amidation of C-terminal carboxyl groups.

A residue, e.g., an amino acid, of a polypeptide of the invention can also be
15 replaced by an amino acid (or peptidomimetic residue) of the opposite chirality. Thus, any amino acid naturally occurring in the L-configuration (which can also be referred to as the R or S, depending upon the structure of the chemical entity) can be replaced with the amino acid of the same chemical structural type or a peptidomimetic, but of the opposite chirality, referred to as the D- amino acid, but also can be referred to as the R- or S- form.

20 The invention also provides methods for modifying the polypeptides of the invention by either natural processes, such as post-translational processing (e.g., phosphorylation, acylation, etc), or by chemical modification techniques, and the resulting modified polypeptides. Modifications can occur anywhere in the polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be
25 appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also a given polypeptide may have many types of modifications. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative,
30 covalent attachment of a phosphatidylinositol, cross-linking cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, and transfer-

RNA mediated addition of amino acids to protein such as arginylation. See, e.g., Creighton, T.E., *Proteins – Structure and Molecular Properties* 2nd Ed., W.H. Freeman and Company, New York (1993); *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York, pp. 1-12 (1983).

5 Solid-phase chemical peptide synthesis methods can also be used to synthesize the polypeptide or fragments of the invention. Such method have been known in the art since the early 1960's (Merrifield, R. B., *J. Am. Chem. Soc.*, 85:2149-2154, 1963) (See also Stewart, J. M. and Young, J. D., *Solid Phase Peptide Synthesis*, 2nd Ed., Pierce Chemical Co., Rockford, Ill., pp. 11-12)) and have recently been employed in commercially available
10 laboratory peptide design and synthesis kits (Cambridge Research Biochemicals). Such commercially available laboratory kits have generally utilized the teachings of H. M. Geysen et al, *Proc. Natl. Acad. Sci., USA*, 81:3998 (1984) and provide for synthesizing peptides upon the tips of a multitude of “rods” or “pins” all of which are connected to a single plate. When such a system is utilized, a plate of rods or pins is inverted and inserted into a second plate of
15 corresponding wells or reservoirs, which contain solutions for attaching or anchoring an appropriate amino acid to the pin's or rod's tips. By repeating such a process step, i.e., inverting and inserting the rod's and pin's tips into appropriate solutions, amino acids are built into desired peptides. In addition, a number of available Fmoc peptide synthesis systems are available. For example, assembly of a polypeptide or fragment can be carried out on a
20 solid support using an Applied Biosystems, Inc. Model 431A™ automated peptide synthesizer. Such equipment provides ready access to the peptides of the invention, either by direct synthesis or by synthesis of a series of fragments that can be coupled using other known techniques.

 The invention includes xylanases of the invention with and without signal.
25 The polypeptide comprising a signal sequence of the invention can be a xylanase of the invention or another xylanase or another enzyme or other polypeptide.

 The invention includes immobilized xylanases, anti-xylanase antibodies and fragments thereof. The invention provides methods for inhibiting xylanase activity, e.g., using dominant negative mutants or anti-xylanase antibodies of the invention. The invention
30 includes heterocomplexes, e.g., fusion proteins, heterodimers, etc., comprising the xylanases of the invention.

 Polypeptides of the invention can have a xylanase activity under various conditions, e.g., extremes in pH and/or temperature, oxidizing agents, and the like. The invention provides methods leading to alternative xylanase preparations with different

catalytic efficiencies and stabilities, e.g., towards temperature, oxidizing agents and changing wash conditions. In one aspect, xylanase variants can be produced using techniques of site-directed mutagenesis and/or random mutagenesis. In one aspect, directed evolution can be used to produce a great variety of xylanase variants with alternative specificities and stability.

5 The proteins of the invention are also useful as research reagents to identify xylanase modulators, e.g., activators or inhibitors of xylanase activity. Briefly, test samples (compounds, broths, extracts, and the like) are added to xylanase assays to determine their ability to inhibit substrate cleavage. Inhibitors identified in this way can be used in industry and research to reduce or prevent undesired proteolysis. As with xylanases, inhibitors can be
10 combined to increase the spectrum of activity.

 The enzymes of the invention are also useful as research reagents to digest proteins or in protein sequencing. For example, the xylanases may be used to break polypeptides into smaller fragments for sequencing using, e.g. an automated sequencer.

 The invention also provides methods of discovering new xylanases using the
15 nucleic acids, polypeptides and antibodies of the invention. In one aspect, phagemid libraries are screened for expression-based discovery of xylanases. In another aspect, lambda phage libraries are screened for expression-based discovery of xylanases. Screening of the phage or phagemid libraries can allow the detection of toxic clones; improved access to substrate; reduced need for engineering a host, by-passing the potential for any bias resulting from mass
20 excision of the library; and, faster growth at low clone densities. Screening of phage or phagemid libraries can be in liquid phase or in solid phase. In one aspect, the invention provides screening in liquid phase. This gives a greater flexibility in assay conditions; additional substrate flexibility; higher sensitivity for weak clones; and ease of automation over solid phase screening.

25 The invention provides screening methods using the proteins and nucleic acids of the invention and robotic automation to enable the execution of many thousands of biocatalytic reactions and screening assays in a short period of time, e.g., per day, as well as ensuring a high level of accuracy and reproducibility (see discussion of arrays, below). As a result, a library of derivative compounds can be produced in a matter of weeks. For further
30 teachings on modification of molecules, including small molecules, see PCT/US94/09174.

 Another aspect of the invention is an isolated or purified polypeptide comprising the sequence of one of Group A nucleic acid sequences and sequences substantially identical thereto, or fragments comprising at least about 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof. As discussed above, such

polypeptides may be obtained by inserting a nucleic acid encoding the polypeptide into a vector such that the coding sequence is operably linked to a sequence capable of driving the expression of the encoded polypeptide in a suitable host cell. For example, the expression vector may comprise a promoter, a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

Another aspect of the invention is polypeptides or fragments thereof which have at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or more than about 95% homology to one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or a fragment comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof. Homology may be determined using any of the programs described above which aligns the polypeptides or fragments being compared and determines the extent of amino acid identity or similarity between them. It will be appreciated that amino acid "homology" includes conservative amino acid substitutions such as those described above.

The polypeptides or fragments having homology to one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or a fragment comprising at least about 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof may be obtained by isolating the nucleic acids encoding them using the techniques described above.

Alternatively, the homologous polypeptides or fragments may be obtained through biochemical enrichment or purification procedures. The sequence of potentially homologous polypeptides or fragments may be determined by xylan hydrolase digestion, gel electrophoresis and/or microsequencing. The sequence of the prospective homologous polypeptide or fragment can be compared to one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or a fragment comprising at least about 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof using any of the programs described above.

Another aspect of the invention is an assay for identifying fragments or variants of Group B amino acid sequences and sequences substantially identical thereto, which retain the enzymatic function of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto. For example the fragments or variants of said polypeptides, may be used to catalyze biochemical reactions, which indicate that the fragment

or variant retains the enzymatic activity of the polypeptides in the Group B amino acid sequences.

The assay for determining if fragments of variants retain the enzymatic activity of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto includes the steps of: contacting the polypeptide fragment or variant with a substrate molecule under conditions which allow the polypeptide fragment or variant to function and detecting either a decrease in the level of substrate or an increase in the level of the specific reaction product of the reaction between the polypeptide and substrate.

The polypeptides of Group B amino acid sequences and sequences substantially identical thereto or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof may be used in a variety of applications. For example, the polypeptides or fragments thereof may be used to catalyze biochemical reactions. In accordance with one aspect of the invention, there is provided a process for utilizing the polypeptides of Group B amino acid sequences and sequences substantially identical thereto or polynucleotides encoding such polypeptides for hydrolyzing glycosidic linkages. In such procedures, a substance containing a glycosidic linkage (*e.g.*, a starch) is contacted with one of the polypeptides of Group B amino acid sequences, or sequences substantially identical thereto under conditions which facilitate the hydrolysis of the glycosidic linkage.

The present invention exploits the unique catalytic properties of enzymes. Whereas the use of biocatalysts (*i.e.*, purified or crude enzymes, non-living or living cells) in chemical transformations normally requires the identification of a particular biocatalyst that reacts with a specific starting compound, the present invention uses selected biocatalysts and reaction conditions that are specific for functional groups that are present in many starting compounds, such as small molecules. Each biocatalyst is specific for one functional group, or several related functional groups and can react with many starting compounds containing this functional group.

The biocatalytic reactions produce a population of derivatives from a single starting compound. These derivatives can be subjected to another round of biocatalytic reactions to produce a second population of derivative compounds. Thousands of variations of the original small molecule or compound can be produced with each iteration of biocatalytic derivatization.

Enzymes react at specific sites of a starting compound without affecting the rest of the molecule, a process which is very difficult to achieve using traditional chemical

methods. This high degree of biocatalytic specificity provides the means to identify a single active compound within the library. The library is characterized by the series of biocatalytic reactions used to produce it, a so called "biosynthetic history". Screening the library for biological activities and tracing the biosynthetic history identifies the specific reaction sequence producing the active compound. The reaction sequence is repeated and the structure of the synthesized compound determined. This mode of identification, unlike other synthesis and screening approaches, does not require immobilization technologies and compounds can be synthesized and tested free in solution using virtually any type of screening assay. It is important to note, that the high degree of specificity of enzyme reactions on functional groups allows for the "tracking" of specific enzymatic reactions that make up the biocatalytically produced library.

Many of the procedural steps are performed using robotic automation enabling the execution of many thousands of biocatalytic reactions and screening assays per day as well as ensuring a high level of accuracy and reproducibility. As a result, a library of derivative compounds can be produced in a matter of weeks which would take years to produce using current chemical methods.

In a particular aspect, the invention provides a method for modifying small molecules, comprising contacting a polypeptide encoded by a polynucleotide described herein or enzymatically active fragments thereof with a small molecule to produce a modified small molecule. A library of modified small molecules is tested to determine if a modified small molecule is present within the library which exhibits a desired activity. A specific biocatalytic reaction which produces the modified small molecule of desired activity is identified by systematically eliminating each of the biocatalytic reactions used to produce a portion of the library and then testing the small molecules produced in the portion of the library for the presence or absence of the modified small molecule with the desired activity. The specific biocatalytic reactions which produce the modified small molecule of desired activity is optionally repeated. The biocatalytic reactions are conducted with a group of biocatalysts that react with distinct structural moieties found within the structure of a small molecule, each biocatalyst is specific for one structural moiety or a group of related structural moieties; and each biocatalyst reacts with many different small molecules which contain the distinct structural moiety.

Xylanase signal sequences, prepro and catalytic domains

The invention provides xylanase signal sequences (e.g., signal peptides (SPs)), prepro domains and catalytic domains (CDs). The SPs, prepro domains and/or CDs of the invention can be isolated or recombinant peptides or can be part of a fusion protein, e.g., as a

5 heterologous domain in a chimeric protein. The invention provides nucleic acids encoding these catalytic domains (CDs), prepro domains and signal sequences (SPs, e.g., a peptide having a sequence comprising/ consisting of amino terminal residues of a polypeptide of the invention). In one aspect, the invention provides a signal sequence comprising a peptide comprising/ consisting of a sequence as set forth in residues 1 to 15, 1 to 16, 1 to 17, 1 to 18,

10 1 to 19, 1 to 20, 1 to 21, 1 to 22, 1 to 23, 1 to 24, 1 to 25, 1 to 26, 1 to 27, 1 to 28, 1 to 28, 1 to 30, 1 to 31, 1 to 32, 1 to 33, 1 to 34, 1 to 35, 1 to 36, 1 to 37, 1 to 38, 1 to 39, 1 to 40, 1 to 41, 1 to 42, 1 to 43, 1 to 44 of a polypeptide of the invention.

In one aspect, the invention provides a signal sequence comprising a peptide comprising/ consisting of a sequence as set forth in Table 4 below. For example, in reading

15 Table 4, the invention provides a signal sequence comprising/ consisting of residues 1 to 23 of SEQ ID NO:102 (encoded by SEQ ID NO:101), a signal sequence comprising/ consisting of residues 1 to 41 of SEQ ID NO:104 (encoded by SEQ ID NO:103), etc.

Table 4: exemplary signal sequences of the invention

SEQ ID NO:	Signal sequence (amino acid positions)
101, 102	1-23
103, 104	1-41
105, 106	1-22
109, 110	1-26
11, 12	1-28
113, 114	1-28
119, 120	1-33
121, 122	1-20
123, 124	1-20
131, 132	1-26
135, 136	1-25
139, 140	1-24
141, 142	1-25
143, 144	1-32
147, 148	1-28
149, 150	1-18
15, 16	1-20
151, 152	1-21
153, 154	1-16
155, 156	1-21

157, 158	1-29
159, 160	1-23
161, 162	1-32
163, 164	1-26
165, 166	1-23
167, 168	1-36
169, 170	1-24
17, 18	1-31
171, 172	1-29
173, 174	1-22
175, 176	1-27
177, 178	1-26
179, 180	1-19
181, 182	1-25
183, 184	1-32
185, 186	1-27
187, 188	1-28
19, 20	1-29
191, 192	1-27
193, 194	1-21
195, 196	1-23
197, 198	1-28
199, 200	1-30
203, 204	1-30
205, 206	1-29
207, 208	1-27
209, 210	1-25
21, 22	1-28
211, 212	1-29
215, 216	1-31
217, 218	1-29
219, 220	1-23
221, 222	1-24
223, 224	1-28
225, 226	1-25
227, 228	1-39
229, 230	1-28
23, 24	1-29
231, 232	1-41
233, 234	1-26
235, 236	1-28
237, 238	1-32
239, 240	1-30
241, 242	1-28
243, 244	1-33
245, 246	1-32
249, 250	1-33
253, 254	1-24
255, 256	1-51
259, 260	1-24
261, 262	1-26
263, 264	1-29

267, 268 1-30
27, 28 1-27
271, 272 1-22
273, 274 1-74
277, 278 1-19
279, 280 1-22
283, 284 1-28
287, 288 1-23
289, 290 1-22
295, 296 1-26
299, 300 1-24
301, 302 1-28
303, 304 1-74
305, 306 1-32
309, 310 1-20
311, 312 1-33
313, 314 1-22
315, 316 1-28
319, 320 1-27
325, 326 1-27
327, 328 1-29
329, 330 1-35
33, 34 1-23
331, 332 1-28
333, 334 1-30
335, 336 1-50
339, 340 1-23
341, 342 1-45
347, 348 1-20
349, 350 1-20
351, 352 1-73
353, 354 1-18
355, 356 1-21
357, 358 1-25
359, 360 1-31
361, 362 1-26
365, 366 1-65
367, 368 1-23
369, 370 1-27
39, 40 1-24
41, 42 1-37
45, 46 1-25
47, 48 1-26
5, 6 1-47
51, 52 1-30
53, 54 1-37
55, 56 1-24
57, 58 1-22
59, 60 1-21
63, 64 1-20
65, 66 1-22
67, 68 1-28

69, 70	1-25
7, 8	1-57
73, 74	1-21
75, 76	1-22
77, 78	1-27
79, 80	1-36
83, 84	1-30
87, 88	1-29
89, 90	1-40
9, 10	1-36
95, 96	1-24
99, 100	1-33

The xylanase signal sequences (SPs) and/or prepro sequences of the invention can be isolated peptides, or, sequences joined to another xylanase or a non-xylanase polypeptide, e.g., as a fusion (chimeric) protein. In one aspect, the invention provides

5 polypeptides comprising xylanase signal sequences of the invention. In one aspect, polypeptides comprising xylanase signal sequences SPs and/or prepro of the invention comprise sequences heterologous to a xylanase of the invention (e.g., a fusion protein comprising an SP and/or prepro of the invention and sequences from another xylanase or a non-xylanase protein). In one aspect, the invention provides xylanases of the invention with

10 heterologous SPs and/or prepro sequences, e.g., sequences with a yeast signal sequence. A xylanase of the invention can comprise a heterologous SP and/or prepro in a vector, e.g., a pPIC series vector (Invitrogen, Carlsbad, CA).

In one aspect, SPs and/or prepro sequences of the invention are identified following identification of novel xylanase polypeptides. The pathways by which proteins are

15 sorted and transported to their proper cellular location are often referred to as protein targeting pathways. One of the most important elements in all of these targeting systems is a short amino acid sequence at the amino terminus of a newly synthesized polypeptide called the signal sequence. This signal sequence directs a protein to its appropriate location in the cell and is removed during transport or when the protein reaches its final destination. Most

20 lysosomal, membrane, or secreted proteins have an amino-terminal signal sequence that marks them for translocation into the lumen of the endoplasmic reticulum. More than 100 signal sequences for proteins in this group have been determined. The signal sequences can vary in length from 13 to 36 amino acid residues. Various methods of recognition of signal sequences are known to those of skill in the art. For example, in one aspect, novel xylanase

25 signal peptides are identified by a method referred to as SignalP. SignalP uses a combined neural network which recognizes both signal peptides and their cleavage sites. (Nielsen, et

al., "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Protein Engineering, vol. 10, no. 1, p. 1-6 (1997).

It should be understood that in some aspects xylanases of the invention may not have SPs and/or prepro sequences, or "domains." In one aspect, the invention provides the xylanases of the invention lacking all or part of an SP and/or a prepro domain. In one aspect, the invention provides a nucleic acid sequence encoding a signal sequence (SP) and/or prepro from one xylanase operably linked to a nucleic acid sequence of a different xylanase or, optionally, a signal sequence (SPs) and/or prepro domain from a non-xylanase protein may be desired.

The invention also provides isolated or recombinant polypeptides comprising signal sequences (SPs), prepro domain and/or catalytic domains (CDs) of the invention and heterologous sequences. The heterologous sequences are sequences not naturally associated (e.g., to a xylanase) with an SP, prepro domain and/or CD. The sequence to which the SP, prepro domain and/or CD are not naturally associated can be on the SP's, prepro domain and/or CD's amino terminal end, carboxy terminal end, and/or on both ends of the SP and/or CD. In one aspect, the invention provides an isolated or recombinant polypeptide comprising (or consisting of) a polypeptide comprising a signal sequence (SP), prepro domain and/or catalytic domain (CD) of the invention with the proviso that it is not associated with any sequence to which it is naturally associated (e.g., a xylanase sequence). Similarly in one aspect, the invention provides isolated or recombinant nucleic acids encoding these polypeptides. Thus, in one aspect, the isolated or recombinant nucleic acid of the invention comprises coding sequence for a signal sequence (SP), prepro domain and/or catalytic domain (CD) of the invention and a heterologous sequence (i.e., a sequence not naturally associated with the a signal sequence (SP), prepro domain and/or catalytic domain (CD) of the invention). The heterologous sequence can be on the 3' terminal end, 5' terminal end, and/or on both ends of the SP, prepro domain and/or CD coding sequence.

Hybrid (chimeric) xylanases and peptide libraries

In one aspect, the invention provides hybrid xylanases and fusion proteins, including peptide libraries, comprising sequences of the invention. The peptide libraries of the invention can be used to isolate peptide modulators (e.g., activators or inhibitors) of targets, such as xylanase substrates, receptors, enzymes. The peptide libraries of the invention can be used to identify formal binding partners of targets, such as ligands, e.g., cytokines, hormones and the like. In one aspect, the invention provides chimeric proteins

comprising a signal sequence (SP), prepro domain and/or catalytic domain (CD) of the invention or a combination thereof and a heterologous sequence (see above).

In one aspect, the fusion proteins of the invention (e.g., the peptide moiety) are conformationally stabilized (relative to linear peptides) to allow a higher binding affinity for targets. The invention provides fusions of xylanases of the invention and other peptides, including known and random peptides. They can be fused in such a manner that the structure of the xylanases is not significantly perturbed and the peptide is metabolically or structurally conformationally stabilized. This allows the creation of a peptide library that is easily monitored both for its presence within cells and its quantity.

Amino acid sequence variants of the invention can be characterized by a predetermined nature of the variation, a feature that sets them apart from a naturally occurring form, e.g., an allelic or interspecies variation of a xylanase sequence. In one aspect, the variants of the invention exhibit the same qualitative biological activity as the naturally occurring analogue. Alternatively, the variants can be selected for having modified characteristics. In one aspect, while the site or region for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed xylanase variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, as discussed herein for example, M13 primer mutagenesis and PCR mutagenesis. Screening of the mutants can be done using, e.g., assays of xylan hydrolysis. In alternative aspects, amino acid substitutions can be single residues; insertions can be on the order of from about 1 to 20 amino acids, although considerably larger insertions can be done. Deletions can range from about 1 to about 20, 30, 40, 50, 60, 70 residues or more. To obtain a final derivative with the optimal properties, substitutions, deletions, insertions or any combination thereof may be used. Generally, these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances.

The invention provides xylanases where the structure of the polypeptide backbone, the secondary or the tertiary structure, e.g., an alpha-helical or beta-sheet structure, has been modified. In one aspect, the charge or hydrophobicity has been modified. In one aspect, the bulk of a side chain has been modified. Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative. For example, substitutions can be made which more significantly affect the structure of the

polypeptide backbone in the area of the alteration, for example a alpha-helical or a beta-sheet structure; a charge or a hydrophobic site of the molecule, which can be at an active site; or a side chain. The invention provides substitutions in polypeptide of the invention where (a) a hydrophilic residues, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g. lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g. glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g. phenylalanine, is substituted for (or by) one not having a side chain, e.g. glycine. The variants can exhibit the same qualitative biological activity (i.e. xylanase activity) although variants can be selected to modify the characteristics of the xylanases as needed.

In one aspect, xylanases of the invention comprise epitopes or purification tags, signal sequences or other fusion sequences, etc. In one aspect, the xylanases of the invention can be fused to a random peptide to form a fusion polypeptide. By "fused" or "operably linked" herein is meant that the random peptide and the xylanase are linked together, in such a manner as to minimize the disruption to the stability of the xylanase structure, e.g., it retains xylanase activity. The fusion polypeptide (or fusion polynucleotide encoding the fusion polypeptide) can comprise further components as well, including multiple peptides at multiple loops.

In one aspect, the peptides and nucleic acids encoding them are randomized, either fully randomized or they are biased in their randomization, e.g. in nucleotide/residue frequency generally or per position. "Randomized" means that each nucleic acid and peptide consists of essentially random nucleotides and amino acids, respectively. In one aspect, the nucleic acids which give rise to the peptides can be chemically synthesized, and thus may incorporate any nucleotide at any position. Thus, when the nucleic acids are expressed to form peptides, any amino acid residue may be incorporated at any position. The synthetic process can be designed to generate randomized nucleic acids, to allow the formation of all or most of the possible combinations over the length of the nucleic acid, thus forming a library of randomized nucleic acids. The library can provide a sufficiently structurally diverse population of randomized expression products to affect a probabilistically sufficient range of cellular responses to provide one or more cells exhibiting a desired response. Thus, the invention provides an interaction library large enough so that at least one of its members will have a structure that gives it affinity for some molecule, protein, or other factor.

Xylanases are multidomain enzymes that consist optionally of a signal peptide, a carbohydrate binding module, a xylanase catalytic domain, a linker and/or another catalytic domain.

The invention provides a means for generating chimeric polypeptides which may encode biologically active hybrid polypeptides (*e.g.*, hybrid xylanases). In one aspect, the original polynucleotides encode biologically active polypeptides. The method of the invention produces new hybrid polypeptides by utilizing cellular processes which integrate the sequence of the original polynucleotides such that the resulting hybrid polynucleotide encodes a polypeptide demonstrating activities derived from the original biologically active polypeptides. For example, the original polynucleotides may encode a particular enzyme from different microorganisms. An enzyme encoded by a first polynucleotide from one organism or variant may, for example, function effectively under a particular environmental condition, *e.g.* high salinity. An enzyme encoded by a second polynucleotide from a different organism or variant may function effectively under a different environmental condition, such as extremely high temperatures. A hybrid polynucleotide containing sequences from the first and second original polynucleotides may encode an enzyme which exhibits characteristics of both enzymes encoded by the original polynucleotides. Thus, the enzyme encoded by the hybrid polynucleotide may function effectively under environmental conditions shared by each of the enzymes encoded by the first and second polynucleotides, *e.g.*, high salinity and extreme temperatures.

Enzymes encoded by the polynucleotides of the invention include, but are not limited to, hydrolases, such as xylanases. Glycosidase hydrolases were first classified into families in 1991, see, *e.g.*, Henrissat (1991) *Biochem. J.* 280:309-316. Since then, the classifications have been continually updated, see, *e.g.*, Henrissat (1993) *Biochem. J.* 293:781-788; Henrissat (1996) *Biochem. J.* 316:695-696; Henrissat (2000) *Plant Physiology* 124:1515-1519. There are 87 identified families of glycosidase hydrolases. In one aspect, the xylanases of the invention may be categorized in families 8, 10, 11, 26 and 30. In one aspect, the invention also provides xylanase-encoding nucleic acids with a common novelty in that they are derived from a common family, *e.g.*, family 5, 6, 8, 10, 11, 26 or 30, as set forth in Table 5, below.

Table 5

SEQ ID	FAMILY
9, 10	8
1, 2	8

5, 6	8
7, 8	8
99, 100	10
11, 12	10
127, 128	10
27, 28	10
97, 98	10
45, 46	10
141, 142	10
107, 108	10
129, 130	10
93, 94	10
63, 64	10
25, 26	10
49, 50	10
67, 68	10
85, 86	10
29, 30	10
51, 52	10
35, 36	10
147, 148	10
119, 120	10
123, 124	10
249, 250	10
149, 150	10
83, 84	10
43, 44	10
133, 134	10
113, 114	10
105, 106	10
75, 76	10
111, 112	10
117, 118	10
115, 116	10
125, 126	10
137, 138	10
135, 136	10
69, 70	10
89, 90	10
31, 32	10
13, 14	10
65, 66	10
57, 58	10
77, 78	10
73, 74	10
109, 110	10
59, 60	10
71, 72	10
139, 140	10
55, 56	10
15, 16	10

131, 132	10
95, 96	10
101, 102	10
39, 40	10
143, 144	10
103, 104	10
17, 18	10
53, 54	10
21, 22	10
151, 152	10
23, 24	10
121, 122	10
41, 42	10
47, 48	10
247, 248	10
33, 34	10
19, 20	10
87, 88	10
81, 82	10
91, 92	10
61, 62	10
37, 38	10
79, 80	10
231, 232	11
157, 158	11
189, 190	11
167, 168	11
207, 208	11
251, 252	11
213, 214	11
177, 178	11
187, 188	11
205, 206	11
211, 212	11
197, 198	11
209, 210	11
185, 186	11
229, 230	11
223, 224	11
179, 180	11
193, 194	11
173, 174	11
217, 218	11
153, 154	11
219, 220	11
183, 184	11
253, 254	11
199, 200	11
255, 256	11
155, 156	11
169, 170	11

195, 196	11
215, 216	11
191, 192	11
175, 176	11
161, 162	11
221, 222	11
225, 226	11
163, 164	11
159, 160	11
233, 234	11
171, 172	11
203, 204	11
181, 182	11
227, 228	11
165, 166	11
257, 258	26
237, 238	30
241, 242	30
239, 240	30
245, 246	30
235, 236	30
313, 314	30
345, 346	10
321, 322	10
323, 324	10
315, 316	10
201, 202	10
265, 266	10
145, 146	10
287, 288	10
293, 294	10
351, 352	10
311, 312	10
279, 280	10
289, 290	10
283, 284	10
373, 374	10
337, 338	10
371, 372	10
291, 292	10
3, 4	10
307, 308	10
343, 344	10
349, 350	10
329, 330	10
355, 356	10
339, 340	10
295, 296	10
333, 334	10
281, 282	10
361, 362	10

347, 348	10
319, 320	10
357, 358	10
365, 366	10
273, 274	10
277, 278	10
271, 272	10
285, 286	10
259, 260	10
325, 326	10
331, 332	10
359, 360	10
303, 304	10
363, 364	10
305, 306	10
341, 342	10
375, 376	11
377, 378	11
379, 380	11
301, 302	11
309, 310	11
263, 264	11
269, 270	11
353, 354	11
299, 300	11
367, 368	11
261, 262	11
369, 370	11
267, 268	11
317, 318	11
297, 298	11
327, 328	5
275, 276	6

A hybrid polypeptide resulting from the method of the invention may exhibit specialized enzyme activity not displayed in the original enzymes. For example, following recombination and/or reductive reassortment of polynucleotides encoding hydrolase

5 activities, the resulting hybrid polypeptide encoded by a hybrid polynucleotide can be screened for specialized hydrolase activities obtained from each of the original enzymes, i.e. the type of bond on which the hydrolase acts and the temperature at which the hydrolase functions. Thus, for example, the hydrolase may be screened to ascertain those chemical functionalities which distinguish the hybrid hydrolase from the original hydrolases, such as:

10 (a) amide (peptide bonds), i.e., xylanases; (b) ester bonds, i.e., esterases and lipases; (c) acetals, i.e., glycosidases and, for example, the temperature, pH or salt concentration at which the hybrid polypeptide functions.

Sources of the original polynucleotides may be isolated from individual organisms ("isolates"), collections of organisms that have been grown in defined media ("enrichment cultures"), or, uncultivated organisms ("environmental samples"). The use of a culture-independent approach to derive polynucleotides encoding novel bioactivities from environmental samples is most preferable since it allows one to access untapped resources of biodiversity.

"Environmental libraries" are generated from environmental samples and represent the collective genomes of naturally occurring organisms archived in cloning vectors that can be propagated in suitable prokaryotic hosts. Because the cloned DNA is initially extracted directly from environmental samples, the libraries are not limited to the small fraction of prokaryotes that can be grown in pure culture. Additionally, a normalization of the environmental DNA present in these samples could allow more equal representation of the DNA from all of the species present in the original sample. This can dramatically increase the efficiency of finding interesting genes from minor constituents of the sample which may be under-represented by several orders of magnitude compared to the dominant species.

For example, gene libraries generated from one or more uncultivated microorganisms are screened for an activity of interest. Potential pathways encoding bioactive molecules of interest are first captured in prokaryotic cells in the form of gene expression libraries. Polynucleotides encoding activities of interest are isolated from such libraries and introduced into a host cell. The host cell is grown under conditions which promote recombination and/or reductive reassortment creating potentially active biomolecules with novel or enhanced activities.

Additionally, subcloning may be performed to further isolate sequences of interest. In subcloning, a portion of DNA is amplified, digested, generally by restriction enzymes, to cut out the desired sequence, the desired sequence is ligated into a recipient vector and is amplified. At each step in subcloning, the portion is examined for the activity of interest, in order to ensure that DNA that encodes the structural protein has not been excluded. The insert may be purified at any step of the subcloning, for example, by gel electrophoresis prior to ligation into a vector or where cells containing the recipient vector and cells not containing the recipient vector are placed on selective media containing, for example, an antibiotic, which will kill the cells not containing the recipient vector. Specific methods of subcloning cDNA inserts into vectors are well-known in the art (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press

(1989)). In another aspect, the enzymes of the invention are subclones. Such subclones may differ from the parent clone by, for example, length, a mutation, a tag or a label.

In one aspect, the signal sequences of the invention are identified following identification of novel xylanase polypeptides. The pathways by which proteins are sorted and transported to their proper cellular location are often referred to as protein targeting pathways. One of the most important elements in all of these targeting systems is a short amino acid sequence at the amino terminus of a newly synthesized polypeptide called the signal sequence. This signal sequence directs a protein to its appropriate location in the cell and is removed during transport or when the protein reaches its final destination. Most lysosomal, membrane, or secreted proteins have an amino-terminal signal sequence that marks them for translocation into the lumen of the endoplasmic reticulum. More than 100 signal sequences for proteins in this group have been determined. The sequences vary in length from 13 to 36 amino acid residues. Various methods of recognition of signal sequences are known to those of skill in the art. In one aspect, the peptides are identified by a method referred to as SignalP. SignalP uses a combined neural network which recognizes both signal peptides and their cleavage sites. See, e.g., Nielsen (1997) "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Protein Engineering, vol. 10, no. 1, p. 1-6. It should be understood that some of the xylanases of the invention may or may not contain signal sequences. It may be desirable to include a nucleic acid sequence encoding a signal sequence from one xylanase operably linked to a nucleic acid sequence of a different xylanase or, optionally, a signal sequence from a non-xylanase protein may be desired.

The microorganisms from which the polynucleotide may be prepared include prokaryotic microorganisms, such as *Eubacteria* and *Archaeobacteria* and lower eukaryotic microorganisms such as fungi, some algae and protozoa. Polynucleotides may be isolated from environmental samples in which case the nucleic acid may be recovered without culturing of an organism or recovered from one or more cultured organisms. In one aspect, such microorganisms may be extremophiles, such as hyperthermophiles, psychrophiles, psychrotrophs, halophiles, barophiles and acidophiles. Polynucleotides encoding enzymes isolated from extremophilic microorganisms can be used. Such enzymes may function at temperatures above 100°C in terrestrial hot springs and deep sea thermal vents, at temperatures below 0°C in arctic waters, in the saturated salt environment of the Dead Sea, at pH values around 0 in coal deposits and geothermal sulfur-rich springs, or at pH values greater than 11 in sewage sludge. For example, several esterases and lipases cloned and

expressed from extremophilic organisms show high activity throughout a wide range of temperatures and pHs.

Polynucleotides selected and isolated as hereinabove described are introduced into a suitable host cell. A suitable host cell is any cell which is capable of promoting recombination and/or reductive reassortment. The selected polynucleotides are preferably already in a vector which includes appropriate control sequences. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or preferably, the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis *et al.*, 1986).

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*, *Streptomyces*, *Salmonella typhimurium*; fungal cells, such as yeast; insect cells such as *Drosophila S2* and *Spodoptera Sf9*; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; and plant cells. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

With particular references to various mammalian cell culture systems that can be employed to express recombinant protein, examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described in "SV40-transformed simian cells support the replication of early SV40 mutants" (Gluzman, 1981) and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

In another aspect, it is envisioned the method of the present invention can be used to generate novel polynucleotides encoding biochemical pathways from one or more operons or gene clusters or portions thereof. For example, bacteria and many eukaryotes have a coordinated mechanism for regulating genes whose products are involved in related processes. The genes are clustered, in structures referred to as "gene clusters," on a single chromosome and are transcribed together under the control of a single regulatory sequence, including a single promoter which initiates transcription of the entire cluster. Thus, a gene cluster is a group of adjacent genes that are either identical or related, usually as to their function. An example of a biochemical pathway encoded by gene clusters are polyketides.

Gene cluster DNA can be isolated from different organisms and ligated into vectors, particularly vectors containing expression regulatory sequences which can control and regulate the production of a detectable protein or protein-related array activity from the ligated gene clusters. Use of vectors which have an exceptionally large capacity for exogenous DNA introduction are particularly appropriate for use with such gene clusters and are described by way of example herein to include the f-factor (or fertility factor) of *E. coli*. This f-factor of *E. coli* is a plasmid which affects high-frequency transfer of itself during conjugation and is ideal to achieve and stably propagate large DNA fragments, such as gene clusters from mixed microbial samples. One aspect of the invention is to use cloning vectors, referred to as "fosmids" or bacterial artificial chromosome (BAC) vectors. These are derived from *E. coli* f-factor which is able to stably integrate large segments of genomic DNA. When integrated with DNA from a mixed uncultured environmental sample, this makes it possible to achieve large genomic fragments in the form of a stable "environmental DNA library." Another type of vector for use in the present invention is a cosmid vector. Cosmid vectors were originally designed to clone and propagate large segments of genomic DNA. Cloning into cosmid vectors is described in detail in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press (1989). Once ligated into an appropriate vector, two or more vectors containing different polyketide synthase gene clusters can be introduced into a suitable host cell. Regions of partial sequence homology shared by the gene clusters will promote processes which result in sequence reorganization resulting in a hybrid gene cluster. The novel hybrid gene cluster can then be screened for enhanced activities not found in the original gene clusters.

Therefore, in a one aspect, the invention relates to a method for producing a biologically active hybrid polypeptide and screening such a polypeptide for enhanced activity by:

- 1) introducing at least a first polynucleotide in operable linkage and a second polynucleotide in operable linkage, the at least first polynucleotide and second polynucleotide sharing at least one region of partial sequence homology, into a suitable host cell;
- 2) growing the host cell under conditions which promote sequence reorganization resulting in a hybrid polynucleotide in operable linkage;
- 3) expressing a hybrid polypeptide encoded by the hybrid polynucleotide;
- 4) screening the hybrid polypeptide under conditions which promote identification of enhanced biological activity; and

- 5) isolating the a polynucleotide encoding the hybrid polypeptide.

Methods for screening for various enzyme activities are known to those of skill in the art and are discussed throughout the present specification. Such methods may be employed when isolating the polypeptides and polynucleotides of the invention.

5 Screening Methodologies and "On-line" Monitoring Devices

In practicing the methods of the invention, a variety of apparatus and methodologies can be used to in conjunction with the polypeptides and nucleic acids of the invention, e.g., to screen polypeptides for xylanase activity (e.g., assays such as hydrolysis of casein in zymograms, the release of fluorescence from gelatin, or the release of p-nitroanalide
10 from various small peptide substrates), to screen compounds as potential modulators, e.g., activators or inhibitors, of a xylanase activity, for antibodies that bind to a polypeptide of the invention, for nucleic acids that hybridize to a nucleic acid of the invention, to screen for cells expressing a polypeptide of the invention and the like. In addition to the array formats described in detail below for screening samples, alternative formats can also be used to
15 practice the methods of the invention. Such formats include, for example, mass spectrometers, chromatographs, e.g., high-throughput HPLC and other forms of liquid chromatography, and smaller formats, such as 1536-well plates, 384-well plates and so on. High throughput screening apparatus can be adapted and used to practice the methods of the invention, see, e.g., U.S. Patent Application No. 20020001809.

20 *Capillary Arrays*

Nucleic acids or polypeptides of the invention can be immobilized to or applied to an array. Arrays can be used to screen for or monitor libraries of compositions (e.g., small molecules, antibodies, nucleic acids, etc.) for their ability to bind to or modulate the activity of a nucleic acid or a polypeptide of the invention. Capillary arrays, such as the
25 GIGAMATRIX™, Diversa Corporation, San Diego, CA; and arrays described in, e.g., U.S. Patent Application No. 20020080350 A1; WO 0231203 A; WO 0244336 A, provide an alternative apparatus for holding and screening samples. In one aspect, the capillary array includes a plurality of capillaries formed into an array of adjacent capillaries, wherein each capillary comprises at least one wall defining a lumen for retaining a sample. The lumen may
30 be cylindrical, square, hexagonal or any other geometric shape so long as the walls form a lumen for retention of a liquid or sample. The capillaries of the capillary array can be held together in close proximity to form a planar structure. The capillaries can be bound together, by being fused (e.g., where the capillaries are made of glass), glued, bonded, or clamped side-

by-side. Additionally, the capillary array can include interstitial material disposed between adjacent capillaries in the array, thereby forming a solid planar device containing a plurality of through-holes.

A capillary array can be formed of any number of individual capillaries, for example, a range from 100 to 4,000,000 capillaries. Further, a capillary array having about 100,000 or more individual capillaries can be formed into the standard size and shape of a Microtiter® plate for fitment into standard laboratory equipment. The lumens are filled manually or automatically using either capillary action or microinjection using a thin needle. Samples of interest may subsequently be removed from individual capillaries for further analysis or characterization. For example, a thin, needle-like probe is positioned in fluid communication with a selected capillary to either add or withdraw material from the lumen.

In a single-pot screening assay, the assay components are mixed yielding a solution of interest, prior to insertion into the capillary array. The lumen is filled by capillary action when at least a portion of the array is immersed into a solution of interest. Chemical or biological reactions and/or activity in each capillary are monitored for detectable events. A detectable event is often referred to as a "hit", which can usually be distinguished from "non-hit" producing capillaries by optical detection. Thus, capillary arrays allow for massively parallel detection of "hits".

In a multi-pot screening assay, a polypeptide or nucleic acid, e.g., a ligand, can be introduced into a first component, which is introduced into at least a portion of a capillary of a capillary array. An air bubble can then be introduced into the capillary behind the first component. A second component can then be introduced into the capillary, wherein the second component is separated from the first component by the air bubble. The first and second components can then be mixed by applying hydrostatic pressure to both sides of the capillary array to collapse the bubble. The capillary array is then monitored for a detectable event resulting from reaction or non-reaction of the two components.

In a binding screening assay, a sample of interest can be introduced as a first liquid labeled with a detectable particle into a capillary of a capillary array, wherein the lumen of the capillary is coated with a binding material for binding the detectable particle to the lumen. The first liquid may then be removed from the capillary tube, wherein the bound detectable particle is maintained within the capillary, and a second liquid may be introduced into the capillary tube. The capillary is then monitored for a detectable event resulting from reaction or non-reaction of the particle with the second liquid.

Arrays, or "Biochips"

Nucleic acids or polypeptides of the invention can be immobilized to or applied to an array. Arrays can be used to screen for or monitor libraries of compositions (e.g., small molecules, antibodies, nucleic acids, etc.) for their ability to bind to or modulate the activity of a nucleic acid or a polypeptide of the invention. For example, in one aspect of the invention, a monitored parameter is transcript expression of a xylanase gene. One or more, or, all the transcripts of a cell can be measured by hybridization of a sample comprising transcripts of the cell, or, nucleic acids representative of or complementary to transcripts of a cell, by hybridization to immobilized nucleic acids on an array, or "biochip." By using an "array" of nucleic acids on a microchip, some or all of the transcripts of a cell can be simultaneously quantified. Alternatively, arrays comprising genomic nucleic acid can also be used to determine the genotype of a newly engineered strain made by the methods of the invention. Polypeptide arrays" can also be used to simultaneously quantify a plurality of proteins. The present invention can be practiced with any known "array," also referred to as a "microarray" or "nucleic acid array" or "polypeptide array" or "antibody array" or "biochip," or variation thereof. Arrays are generically a plurality of "spots" or "target elements," each target element comprising a defined amount of one or more biological molecules, e.g., oligonucleotides, immobilized onto a defined area of a substrate surface for specific binding to a sample molecule, e.g., mRNA transcripts.

In practicing the methods of the invention, any known array and/or method of making and using arrays can be incorporated in whole or in part, or variations thereof, as described, for example, in U.S. Patent Nos. 6,277,628; 6,277,489; 6,261,776; 6,258,606; 6,054,270; 6,048,695; 6,045,996; 6,022,963; 6,013,440; 5,965,452; 5,959,098; 5,856,174; 5,830,645; 5,770,456; 5,632,957; 5,556,752; 5,143,854; 5,807,522; 5,800,992; 5,744,305; 5,700,637; 5,556,752; 5,434,049; see also, e.g., WO 99/51773; WO 99/09217; WO 97/46313; WO 96/17958; see also, e.g., Johnston (1998) *Curr. Biol.* 8:R171-R174; Schummer (1997) *Biotechniques* 23:1087-1092; Kern (1997) *Biotechniques* 23:120-124; Solinas-Toldo (1997) *Genes, Chromosomes & Cancer* 20:399-407; Bowtell (1999) *Nature Genetics Supp.* 21:25-32. See also published U.S. patent applications Nos. 20010018642; 20010019827; 20010016322; 20010014449; 20010014448; 20010012537; 20010008765.

Antibodies and Antibody-based screening methods

The invention provides isolated or recombinant antibodies that specifically bind to a xylanase of the invention. These antibodies can be used to isolate, identify or

quantify the xylanases of the invention or related polypeptides. These antibodies can be used to isolate other polypeptides within the scope the invention or other related xylanases. The antibodies can be designed to bind to an active site of a xylanase. Thus, the invention provides methods of inhibiting xylanases using the antibodies of the invention (see discussion
5 above regarding applications for anti-xylanase compositions of the invention).

The invention provides fragments of the enzymes of the invention, including immunogenic fragments of a polypeptide of the invention. The invention provides compositions comprising a polypeptide or peptide of the invention and adjuvants or carriers and the like.

10 The antibodies can be used in immunoprecipitation, staining, immunoaffinity columns, and the like. If desired, nucleic acid sequences encoding for specific antigens can be generated by immunization followed by isolation of polypeptide or nucleic acid, amplification or cloning and immobilization of polypeptide onto an array of the invention. Alternatively, the methods of the invention can be used to modify the structure of an antibody
15 produced by a cell to be modified, e.g., an antibody's affinity can be increased or decreased. Furthermore, the ability to make or modify antibodies can be a phenotype engineered into a cell by the methods of the invention.

Methods of immunization, producing and isolating antibodies (polyclonal and monoclonal) are known to those of skill in the art and described in the scientific and patent
20 literature, see, e.g., Coligan, CURRENT PROTOCOLS IN IMMUNOLOGY, Wiley/Greene, NY (1991); Stites (eds.) BASIC AND CLINICAL IMMUNOLOGY (7th ed.) Lange Medical Publications, Los Altos, CA ("Stites"); Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE (2d ed.) Academic Press, New York, NY (1986); Kohler (1975) Nature 256:495; Harlow (1988) ANTIBODIES, A LABORATORY MANUAL, Cold
25 Spring Harbor Publications, New York. Antibodies also can be generated *in vitro*, e.g., using recombinant antibody binding site expressing phage display libraries, in addition to the traditional *in vivo* methods using animals. See, e.g., Hoogenboom (1997) Trends Biotechnol. 15:62-70; Katz (1997) Annu. Rev. Biophys. Biomol. Struct. 26:27-45.

The polypeptides of Group B amino acid sequences and sequences
30 substantially identical thereto or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof, may also be used to generate antibodies which bind specifically to the polypeptides or fragments. The resulting antibodies may be used in immunoaffinity chromatography procedures to isolate or purify the polypeptide or to determine whether the polypeptide is present in a biological sample. In such procedures, a

protein preparation, such as an extract, or a biological sample is contacted with an antibody capable of specifically binding to one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof.

5 In immunoaffinity procedures, the antibody is attached to a solid support, such as a bead or other column matrix. The protein preparation is placed in contact with the antibody under conditions in which the antibody specifically binds to one of the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragment thereof. After a wash to remove non-specifically bound proteins, the specifically bound
10 polypeptides are eluted.

The ability of proteins in a biological sample to bind to the antibody may be determined using any of a variety of procedures familiar to those skilled in the art. For example, binding may be determined by labeling the antibody with a detectable label such as a fluorescent agent, an enzymatic label, or a radioisotope. Alternatively, binding of the
15 antibody to the sample may be detected using a secondary antibody having such a detectable label thereon. Particular assays include ELISA assays, sandwich assays, radioimmunoassays and Western Blots.

Polyclonal antibodies generated against the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least
20 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, for example, a nonhuman. The antibody so obtained will then bind the polypeptide itself. In this manner, even a sequence encoding only a fragment of the polypeptide can be used to generate antibodies which may bind to the whole native
25 polypeptide. Such antibodies can then be used to isolate the polypeptide from cells expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, *Nature*, 256:495-497, 1975), the trioma
30 technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* 4:72, 1983) and the EBV-hybridoma technique (Cole, *et al.*, 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to the polypeptides

of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof. Alternatively, transgenic mice may be used to express humanized antibodies to these polypeptides or fragments thereof.

5 Antibodies generated against the polypeptides of Group B amino acid sequences and sequences substantially identical thereto, or fragments comprising at least 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, or 150 consecutive amino acids thereof may be used in screening for similar polypeptides from other organisms and samples. In such techniques, polypeptides from the organism are contacted with the antibody and those polypeptides which
10 specifically bind the antibody are detected. Any of the procedures described above may be used to detect antibody binding. One such screening assay is described in "Methods for Measuring Cellulase Activities", *Methods in Enzymology*, Vol 160, pp. 87-116.

Kits

15 The invention provides kits comprising the compositions, e.g., nucleic acids, expression cassettes, vectors, cells, transgenic seeds or plants or plant parts, polypeptides (e.g., xylanases) and/or antibodies of the invention. The kits also can contain instructional material teaching the methodologies and industrial uses of the invention, as described herein.

Whole cell engineering and measuring metabolic parameters

20 The methods of the invention provide whole cell evolution, or whole cell engineering, of a cell to develop a new cell strain having a new phenotype, e.g., a new or modified xylanase activity, by modifying the genetic composition of the cell. The genetic composition can be modified by addition to the cell of a nucleic acid of the invention, e.g., a coding sequence for an enzyme of the invention. See, e.g., WO0229032; WO0196551.

25 To detect the new phenotype, at least one metabolic parameter of a modified cell is monitored in the cell in a "real time" or "on-line" time frame. In one aspect, a plurality of cells, such as a cell culture, is monitored in "real time" or "on-line." In one aspect, a plurality of metabolic parameters is monitored in "real time" or "on-line." Metabolic parameters can be monitored using the xylanases of the invention.

30 Metabolic flux analysis (MFA) is based on a known biochemistry framework. A linearly independent metabolic matrix is constructed based on the law of mass conservation and on the pseudo-steady state hypothesis (PSSH) on the intracellular metabolites. In practicing the methods of the invention, metabolic networks are established, including the:

• identity of all pathway substrates, products and intermediary metabolites
• identity of all the chemical reactions interconverting the pathway metabolites, the stoichiometry of the pathway reactions,
• identity of all the enzymes catalyzing the reactions, the enzyme reaction kinetics,
5 • the regulatory interactions between pathway components, e.g. allosteric interactions, enzyme-enzyme interactions etc,
• intracellular compartmentalization of enzymes or any other supramolecular organization of the enzymes, and,
• the presence of any concentration gradients of metabolites, enzymes or effector
10 molecules or diffusion barriers to their movement.

Once the metabolic network for a given strain is built, mathematic presentation by matrix notion can be introduced to estimate the intracellular metabolic fluxes if the on-line metabolome data is available. Metabolic phenotype relies on the changes of the whole metabolic network within a cell. Metabolic phenotype relies on the change of pathway
15 utilization with respect to environmental conditions, genetic regulation, developmental state and the genotype, etc. In one aspect of the methods of the invention, after the on-line MFA calculation, the dynamic behavior of the cells, their phenotype and other properties are analyzed by investigating the pathway utilization. For example, if the glucose supply is increased and the oxygen decreased during the yeast fermentation, the utilization of
20 respiratory pathways will be reduced and/or stopped, and the utilization of the fermentative pathways will dominate. Control of physiological state of cell cultures will become possible after the pathway analysis. The methods of the invention can help determine how to manipulate the fermentation by determining how to change the substrate supply, temperature, use of inducers, etc. to control the physiological state of cells to move along desirable
25 direction. In practicing the methods of the invention, the MFA results can also be compared with transcriptome and proteome data to design experiments and protocols for metabolic engineering or gene shuffling, etc.

In practicing the methods of the invention, any modified or new phenotype can be conferred and detected, including new or improved characteristics in the cell. Any
30 aspect of metabolism or growth can be monitored.

Monitoring expression of an mRNA transcript

In one aspect of the invention, the engineered phenotype comprises increasing or decreasing the expression of an mRNA transcript (e.g., a xylanase message) or generating

new (e.g., xylanase) transcripts in a cell. This increased or decreased expression can be traced by testing for the presence of a xylanase of the invention or by xylanase activity assays. mRNA transcripts, or messages, also can be detected and quantified by any method known in the art, including, e.g., Northern blots, quantitative amplification reactions, hybridization to arrays, and the like. Quantitative amplification reactions include, e.g., quantitative PCR, including, e.g., quantitative reverse transcription polymerase chain reaction, or RT-PCR; quantitative real time RT-PCR, or "real-time kinetic RT-PCR" (see, e.g., Kreuzer (2001) Br. J. Haematol. 114:313-318; Xia (2001) Transplantation 72:907-914).

In one aspect of the invention, the engineered phenotype is generated by knocking out expression of a homologous gene. The gene's coding sequence or one or more transcriptional control elements can be knocked out, e.g., promoters or enhancers. Thus, the expression of a transcript can be completely ablated or only decreased.

In one aspect of the invention, the engineered phenotype comprises increasing the expression of a homologous gene. This can be effected by knocking out of a negative control element, including a transcriptional regulatory element acting in cis- or trans- , or, mutagenizing a positive control element. One or more, or, all the transcripts of a cell can be measured by hybridization of a sample comprising transcripts of the cell, or, nucleic acids representative of or complementary to transcripts of a cell, by hybridization to immobilized nucleic acids on an array.

Monitoring expression of a polypeptides, peptides and amino acids

In one aspect of the invention, the engineered phenotype comprises increasing or decreasing the expression of a polypeptide (e.g., a xylanase) or generating new polypeptides in a cell. This increased or decreased expression can be traced by determining the amount of xylanase present or by xylanase activity assays. Polypeptides, peptides and amino acids also can be detected and quantified by any method known in the art, including, e.g., nuclear magnetic resonance (NMR), spectrophotometry, radiography (protein radiolabeling), electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, various immunological methods, e.g. immunoprecipitation, immunodiffusion, immunoelectrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), immuno-fluorescent assays, gel electrophoresis (e.g., SDS-PAGE), staining with antibodies, fluorescent activated cell sorter (FACS), pyrolysis mass spectrometry, Fourier-Transform Infrared Spectrometry, Raman spectrometry, GC-MS, and LC-Electrospray and

cap-LC-tandem-electrospray mass spectrometries, and the like. Novel bioactivities can also be screened using methods, or variations thereof, described in U.S. Patent No. 6,057,103. Furthermore, as discussed below in detail, one or more, or, all the polypeptides of a cell can be measured using a protein array.

5 Industrial Applications

The xylanase enzymes of the invention can be highly selective catalysts. They can catalyze reactions with exquisite stereo-, regio- and chemo- selectivities that are unparalleled in conventional synthetic chemistry. Moreover, enzymes are remarkably versatile. The xylanase enzymes of the invention can be tailored to function in organic
10 solvents, operate at extreme pHs (for example, high pHs and low pHs) extreme temperatures (for example, high temperatures and low temperatures), extreme salinity levels (for example, high salinity and low salinity) and catalyze reactions with compounds that are structurally unrelated to their natural, physiological substrates.

Detergent Compositions

15 The invention provides detergent compositions comprising one or more polypeptides (e.g., xylanases) of the invention, and methods of making and using these compositions. The invention incorporates all methods of making and using detergent compositions, see, e.g., U.S. Patent No. 6,413,928; 6,399,561; 6,365,561; 6,380,147. The detergent compositions can be a one and two part aqueous composition, a non-aqueous liquid
20 composition, a cast solid, a granular form, a particulate form, a compressed tablet, a gel and/or a paste and a slurry form. The xylanases of the invention can also be used as a detergent additive product in a solid or a liquid form. Such additive products are intended to supplement or boost the performance of conventional detergent compositions and can be added at any stage of the cleaning process.

25 The actual active enzyme content depends upon the method of manufacture of a detergent composition and is not critical, assuming the detergent solution has the desired enzymatic activity. In one aspect, the amount of xylanase present in the final solution ranges from about 0.001 mg to 0.5 mg per gram of the detergent composition. The particular enzyme chosen for use in the process and products of this invention depends upon the
30 conditions of final utility, including the physical product form, use pH, use temperature, and soil types to be degraded or altered. The enzyme can be chosen to provide optimum activity and stability for any given set of utility conditions. In one aspect, the xylanases of the present invention are active in the pH ranges of from about 4 to about 12 and in the temperature

range of from about 20°C to about 95°C. The detergents of the invention can comprise cationic, semi-polar nonionic or zwitterionic surfactants; or, mixtures thereof.

Xylanases of the invention can be formulated into powdered and liquid detergents having pH between 4.0 and 12.0 at levels of about 0.01 to about 5% (preferably 0.1% to 0.5%) by weight. These detergent compositions can also include other enzymes such as xylanases, cellulases, lipases or endoglycosidases, endo-beta.-1,4-glucanases, beta-glucanases, endo-beta-1,3(4)-glucanases, cutinases, peroxidases, laccases, amylases, glucoamylases, pectinases, reductases, oxidases, phenoloxidases, ligninases, pullulanases, arabinanases, hemicellulases, mannanases, xyloglucanases, xylanases, pectin acetyl esterases, rhamnogalacturonan acetyl esterases, polygalacturonases, rhamnogalacturonases, galactanases, pectin lyases, pectin methylesterases, cellobiohydrolases and/or transglutaminases. These detergent compositions can also include builders and stabilizers.

The addition of xylanases of the invention to conventional cleaning compositions does not create any special use limitation. In other words, any temperature and pH suitable for the detergent is also suitable for the compositions of the invention as long as the enzyme is active at or tolerant of the pH and/or temperature of the intended use. In addition, the xylanases of the invention can be used in a cleaning composition without detergents, again either alone or in combination with builders and stabilizers.

The present invention provides cleaning compositions including detergent compositions for cleaning hard surfaces, detergent compositions for cleaning fabrics, dishwashing compositions, oral cleaning compositions, denture cleaning compositions, and contact lens cleaning solutions.

In one aspect, the invention provides a method for washing an object comprising contacting the object with a polypeptide of the invention under conditions sufficient for washing. A xylanase of the invention may be included as a detergent additive. The detergent composition of the invention may, for example, be formulated as a hand or machine laundry detergent composition comprising a polypeptide of the invention. A laundry additive suitable for pre-treatment of stained fabrics can comprise a polypeptide of the invention. A fabric softener composition can comprise a xylanase of the invention. Alternatively, a xylanase of the invention can be formulated as a detergent composition for use in general household hard surface cleaning operations. In alternative aspects, detergent additives and detergent compositions of the invention may comprise one or more other enzymes such as a xylanase, a lipase, a cutinase, another xylanase, a carbohydrase, a cellulase, a pectinase, a mannanase, an arabinase, a galactanase, a xylanase, an oxidase, e.g.,

a lactase, and/or a peroxidase (see also, above). The properties of the enzyme(s) of the invention are chosen to be compatible with the selected detergent (i.e. pH-optimum, compatibility with other enzymatic and non-enzymatic ingredients, etc.) and the enzyme(s) is present in effective amounts. In one aspect, xylanase enzymes of the invention are used to
5 remove malodorous materials from fabrics. Various detergent compositions and methods for making them that can be used in practicing the invention are described in, e.g., U.S. Patent Nos. 6,333,301; 6,329,333; 6,326,341; 6,297,038; 6,309,871; 6,204,232; 6,197,070; 5,856,164.

When formulated as compositions suitable for use in a laundry machine
10 washing method, the xylanases of the invention can comprise both a surfactant and a builder compound. They can additionally comprise one or more detergent components, e.g., organic polymeric compounds, bleaching agents, additional enzymes, suds suppressors, dispersants, lime-soap dispersants, soil suspension and anti-redeposition agents and corrosion inhibitors. Laundry compositions of the invention can also contain softening agents, as additional
15 detergent components. Such compositions containing carbohydrase can provide fabric cleaning, stain removal, whiteness maintenance, softening, color appearance, dye transfer inhibition and sanitization when formulated as laundry detergent compositions.

The density of the laundry detergent compositions of the invention can range from about 200 to 1500 g/liter, or, about 400 to 1200 g/liter, or, about 500 to 950 g/liter, or,
20 600 to 800 g/liter, of composition; this can be measured at about 20°C.

The "compact" form of laundry detergent compositions of the invention is best reflected by density and, in terms of composition, by the amount of inorganic filler salt. Inorganic filler salts are conventional ingredients of detergent compositions in powder form. In conventional detergent compositions, the filler salts are present in substantial amounts,
25 typically 17% to 35% by weight of the total composition. In one aspect of the compact compositions, the filler salt is present in amounts not exceeding 15% of the total composition, or, not exceeding 10%, or, not exceeding 5% by weight of the composition. The inorganic filler salts can be selected from the alkali and alkaline-earth-metal salts of sulphates and chlorides, e.g., sodium sulphate.

30 Liquid detergent compositions of the invention can also be in a "concentrated form." In one aspect, the liquid detergent compositions can contain a lower amount of water, compared to conventional liquid detergents. In alternative aspects, the water content of the concentrated liquid detergent is less than 40%, or, less than 30%, or, less than 20% by weight

of the detergent composition. Detergent compounds of the invention can comprise formulations as described in WO 97/01629.

Xylanases of the invention can be useful in formulating various cleaning compositions. A number of known compounds are suitable surfactants including nonionic, anionic, cationic, or zwitterionic detergents, can be used, e.g., as disclosed in U.S. Patent Nos. 4,404,128; 4,261,868; 5,204,015. In addition, xylanases can be used, for example, in bar or liquid soap applications, dish care formulations, contact lens cleaning solutions or products, peptide hydrolysis, waste treatment, textile applications, as fusion-cleavage enzymes in protein production, and the like. Xylanases may provide enhanced performance in a detergent composition as compared to another detergent xylanase, that is, the enzyme group may increase cleaning of certain enzyme sensitive stains such as grass or blood, as determined by usual evaluation after a standard wash cycle. Xylanases can be formulated into known powdered and liquid detergents having pH between 6.5 and 12.0 at levels of about 0.01 to about 5% (for example, about 0.1% to 0.5%) by weight. These detergent cleaning compositions can also include other enzymes such as known xylanases, xylanases, amylases, cellulases, lipases or endoglycosidases, as well as builders and stabilizers.

In one aspect, the invention provides detergent compositions having xylanase activity (a xylanase of the invention) for use with fruit, vegetables and/or mud and clay compounds (see, for example, U.S. Pat. No. 5,786,316).

Treating fibers and textiles

The invention provides methods of treating fibers and fabrics using one or more xylanases of the invention. The xylanases can be used in any fiber- or fabric-treating method, which are well known in the art, see, e.g., U.S. Patent No. 6,261,828; 6,077,316; 6,024,766; 6,021,536; 6,017,751; 5,980,581; US Patent Publication No. 20020142438 A1. For example, xylanases of the invention can be used in fiber and/or fabric desizing. In one aspect, the feel and appearance of a fabric is improved by a method comprising contacting the fabric with a xylanase of the invention in a solution. In one aspect, the fabric is treated with the solution under pressure. For example, xylanases of the invention can be used in the removal of stains.

The xylanases of the invention can be used to treat any cellulosic material, including fibers (e.g., fibers from cotton, hemp, flax or linen), sewn and unsewn fabrics, e.g., knits, wovens, denims, yarns, and toweling, made from cotton, cotton blends or natural or manmade cellulose (e.g. originating from xylan-containing cellulose fibers such as from

wood pulp) or blends thereof. Examples of blends are blends of cotton or rayon/viscose with one or more companion material such as wool, synthetic fibers (e.g. polyamide fibers, acrylic fibers, polyester fibers, polyvinyl alcohol fibers, polyvinyl chloride fibers, polyvinylidene chloride fibers, polyurethane fibers, polyurea fibers, aramid fibers), and cellulose-containing
5 fibers (e.g. rayon/viscose, ramie, hemp, flax/linen, jute, cellulose acetate fibers, lyocell).

The textile treating processes of the invention (using xylanases of the invention) can be used in conjunction with other textile treatments, e.g., scouring and bleaching. Scouring is the removal of non-cellulosic material from the cotton fiber, e.g., the cuticle (mainly consisting of waxes) and primary cell wall (mainly consisting of pectin,
10 protein and xyloglucan). A proper wax removal is necessary for obtaining a high wettability. This is needed for dyeing. Removal of the primary cell walls by the processes of the invention improves wax removal and ensures a more even dyeing. Treating textiles with the processes of the invention can improve whiteness in the bleaching process. The main chemical used in scouring is sodium hydroxide in high concentrations and at high
15 temperatures. Bleaching comprises oxidizing the textile. Bleaching typically involves use of hydrogen peroxide as the oxidizing agent in order to obtain either a fully bleached (white) fabric or to ensure a clean shade of the dye.

The invention also provides alkaline xylanases (xylanases active under alkaline conditions). These have wide-ranging applications in textile processing, degumming
20 of plant fibers (e.g., plant bast fibers), treatment of pectic wastewaters, paper-making, and coffee and tea fermentations. See, e.g., Hoondal (2002) Applied Microbiology and Biotechnology 59:409-418.

Treating foods and food processing

The xylanases of the invention have numerous applications in food processing
25 industry. For example, in one aspect, the xylanases of the invention are used to improve the extraction of oil from oil-rich plant material, e.g., oil-rich seeds, for example, soybean oil from soybeans, olive oil from olives, rapeseed oil from rapeseed and/or sunflower oil from sunflower seeds.

The xylanases of the invention can be used for separation of components of
30 plant cell materials. For example, xylanases of the invention can be used in the separation of xylan-rich material (e.g., plant cells) into components. In one aspect, xylanases of the invention can be used to separate xylan-rich or oil-rich crops into valuable protein and oil and hull fractions. The separation process may be performed by use of methods known in the art.

The xylanases of the invention can be used in the preparation of fruit or vegetable juices, syrups, extracts and the like to increase yield. The xylanases of the invention can be used in the enzymatic treatment (e.g., hydrolysis of xylan-comprising plant materials) of various plant cell wall-derived materials or waste materials, e.g. from cereals, grains, wine or juice production, or agricultural residues such as vegetable hulls, bean hulls, sugar beet pulp, olive pulp, potato pulp, and the like. The xylanases of the invention can be used to modify the consistency and appearance of processed fruit or vegetables. The xylanases of the invention can be used to treat plant material to facilitate processing of plant material, including foods, facilitate purification or extraction of plant components. The xylanases of the invention can be used to improve feed value, decrease the water binding capacity, improve the degradability in waste water plants and/or improve the conversion of plant material to ensilage, and the like.

In one aspect, xylanases of the invention are used in baking applications, e.g., cookies and crackers, to hydrolyze arabinoxylans and create non-sticky doughs that are not difficult to machine and to reduce biscuit size. Use xylanases of the invention to hydrolyze arabinoxylans is used to prevent rapid rehydration of the baked product resulting in loss of crispiness and reduced shelf-life. In one aspect, xylanases of the invention are used as additives in dough processing. In one aspect, xylanases of the invention are used in dough conditioning, wherein in one aspect the xylanases possess high activity over a temperature range of about 25-35°C and at near neutral pH (7.0 – 7.5). In one aspect, dough conditioning enzymes can be inactivated at the extreme temperatures of baking (>500°F).

In one aspect, xylanases of the invention are used as additives in dough processing to perform optimally under dough pH and temperature conditions. In one aspect, an enzyme of the invention is used for dough conditioning. In one aspect, a xylanase of the invention possesses high activity over a temperature range of 25-35°C and at near neutral pH (7.0 – 7.5). In one aspect, the enzyme is inactivated at the extreme temperatures of baking, for example, >500°F.

Paper or pulp treatment

The xylanases of the invention can be in paper or pulp treatment or paper deinking. For example, in one aspect, the invention provides a paper treatment process using a xylanase of the invention. In one aspect, the xylanase of the invention is applicable both in reduction of the need for a chemical bleaching agent, such as chlorine dioxide, and in high alkaline and high temperature environments. In one aspect, the xylanase of the invention is a

thermostable alkaline endoxylanase which can effect a greater than 25% reduction in the chlorine dioxide requirement of kraft pulp with a less than 0.5% pulp yield loss. In one aspect, boundary parameters are pH 10, 65-85°C and treatment time of less than 60 minutes at an enzyme loading of less than 0.001 wt%. A pool of xylanases may be tested for the ability to hydrolyze dye-labeled xylan at, for example, pH 10 and 60°C. The enzymes that test positive under these conditions may then be evaluated at, for example pH 10 and 70°C. Alternatively, enzymes may be tested at pH 8 and pH 10 at 70°C. In discovery of xylanases desirable in the pulp and paper industry libraries from high temperature or highly alkaline environments were targeted. Specifically, these libraries were screened for enzymes functioning at alkaline pH and a temperature of approximately 45°C. In another aspect, the xylanases of the invention are useful in the pulp and paper industry in degradation of a lignin hemicellulose linkage, in order to release the lignin.

Animal feeds and food or feed additives

The invention provides methods for treating animal feeds and foods and food or feed additives using xylanases of the invention, animals including mammals (e.g., humans), birds, fish and the like. The invention provides animal feeds, foods, and additives comprising xylanases of the invention. In one aspect, treating animal feeds, foods and additives using xylanases of the invention can help in the availability of nutrients, e.g., starch, protein, and the like, in the animal feed or additive. By breaking down difficult to digest proteins or indirectly or directly unmasking starch (or other nutrients), the xylanase makes nutrients more accessible to other endogenous or exogenous enzymes. The xylanase can also simply cause the release of readily digestible and easily absorbed nutrients and sugars.

When added to animal feed, xylanases of the invention improve the *in vivo* break-down of plant cell wall material partly due to a reduction of the intestinal viscosity (see, e.g., Bedford et al., Proceedings of the 1st Symposium on Enzymes in Animal Nutrition, 1993, pp. 73-77), whereby a better utilization of the plant nutrients by the animal is achieved. Thus, by using xylanases of the invention in feeds the growth rate and/or feed conversion ratio (i.e. the weight of ingested feed relative to weight gain) of the animal is improved.

The animal feed additive of the invention may be a granulated enzyme product which may readily be-mixed with feed components. Alternatively, feed additives of the invention can form a component of a pre-mix. The granulated enzyme product of the invention may be coated or uncoated. The particle size of the enzyme granulates can be compatible with that of feed and pre-mix components. This provides a safe and convenient

mean of incorporating enzymes into feeds. Alternatively, the animal feed additive of the invention may be a stabilized liquid composition. This may be an aqueous or oil-based slurry. See, e.g., U.S. Patent No. 6,245,546.

Xylanases of the present invention, in the modification of animal feed or a food, can process the food or feed either *in vitro* (by modifying components of the feed or food) or *in vivo*. Xylanases can be added to animal feed or food compositions containing high amounts of xylans, e.g. feed or food containing plant material from cereals, grains and the like. When added to the feed or food the xylanase significantly improves the *in vivo* break-down of xylan-containing material, e.g., plant cell walls, whereby a better utilization of the plant nutrients by the animal (e.g., human) is achieved. In one aspect, the growth rate and/or feed conversion ratio (i.e. the weight of ingested feed relative to weight gain) of the animal is improved. For example a partially or indigestible xylan-comprising protein is fully or partially degraded by a xylanase of the invention, e.g. in combination with another enzyme, e.g., beta-galactosidase, to peptides and galactose and/or galactooligomers. These enzyme digestion products are more digestible by the animal. Thus, xylanases of the invention can contribute to the available energy of the feed or food. Also, by contributing to the degradation of xylan-comprising proteins, a xylanase of the invention can improve the digestibility and uptake of carbohydrate and non-carbohydrate feed or food constituents such as protein, fat and minerals.

In another aspect, xylanase of the invention can be supplied by expressing the enzymes directly in transgenic feed crops (as, e.g., transgenic plants, seeds and the like), such as grains, cereals, corn, soy bean, rape seed, lupin and the like. As discussed above, the invention provides transgenic plants, plant parts and plant cells comprising a nucleic acid sequence encoding a polypeptide of the invention. In one aspect, the nucleic acid is expressed such that the xylanase of the invention is produced in recoverable quantities. The xylanase can be recovered from any plant or plant part. Alternatively, the plant or plant part containing the recombinant polypeptide can be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

In one aspect, the invention provides methods for removing oligosaccharides from feed prior to consumption by an animal subject using a xylanase of the invention. In this process a feed is formed having an increased metabolizable energy value. In addition to xylanases of the invention, galactosidases, cellulases and combinations thereof can be used. In one aspect, the enzyme is added in an amount equal to between about 0.1% and 1% by

weight of the feed material. In one aspect, the feed is a cereal, a wheat, a grain, a soybean (e.g., a ground soybean) material. See, e.g., U.S. Patent No. 6,399,123.

In another aspect, the invention provides methods for utilizing xylanase as a nutritional supplement in the diets of animals by preparing a nutritional supplement
5 containing a recombinant xylanase enzyme comprising at least thirty contiguous amino acids of an amino acid of Group B amino acid sequences, and administering the nutritional supplement to an animal to increase the utilization of xylan contained in food ingested by the animal.

In yet another aspect, the invention provides an edible pelletized enzyme
10 delivery matrix and method of use for delivery of xylanase to an animal, for example as a nutritional supplement. The enzyme delivery matrix readily releases a xylanase enzyme, such as one having an amino acid sequence of group B amino acid sequences, or at least 30 contiguous amino acids thereof, in aqueous media, such as, for example, the digestive fluid of an animal. The invention enzyme delivery matrix is prepared from a granulate edible carrier
15 selected from such components as grain germ that is spent of oil, hay, alfalfa, timothy, soy hull, sunflower seed meal, wheat midd, and the like, that readily disperse the recombinant enzyme contained therein into aqueous media. In use, the edible pelletized enzyme delivery matrix is administered to an animal to delivery of xylanase to the animal. Suitable grain-based substrates may comprise or be derived from any suitable edible grain, such as wheat,
20 corn, soy, sorghum, alfalfa, barley, and the like. An exemplary grain-based substrate is a corn-based substrate. The substrate may be derived from any suitable part of the grain, but is preferably a grain germ approved for animal feed use, such as corn germ that is obtained in a wet or dry milling process. The grain germ preferably comprises spent germ, which is grain germ from which oil has been expelled, such as by pressing or hexane or other solvent
25 extraction. Alternatively, the grain germ is expeller extracted, that is, the oil has been removed by pressing.

The enzyme delivery matrix of the invention is in the form of discrete plural particles, pellets or granules. By "granules" is meant particles that are compressed or compacted, such as by a pelletizing, extrusion, or similar compacting to remove water from
30 the matrix. Such compression or compacting of the particles also promotes intraparticle cohesion of the particles. For example, the granules can be prepared by pelletizing the grain-based substrate in a pellet mill. The pellets prepared thereby are ground or crumbled to a granule size suitable for use as an adjuvant in animal feed. Since the matrix is itself approved for use in animal feed, it can be used as a diluent for delivery of enzymes in animal feed.

Preferably, the enzyme delivery matrix is in the form of granules having a granule size ranging from about 4 to about 400 mesh (USS); more preferably, about 8 to about 80 mesh; and most preferably about 14 to about 20 mesh. If the grain germ is spent via solvent extraction, use of a lubricity agent such as corn oil may be necessary in the pelletizer, but such a lubricity agent ordinarily is not necessary if the germ is expeller extracted. In other aspects of the invention, the matrix is prepared by other compacting or compressing processes such as, for example, by extrusion of the grain-based substrate through a die and grinding of the extrudate to a suitable granule size.

The enzyme delivery matrix may further include a polysaccharide component as a cohesiveness agent to enhance the cohesiveness of the matrix granules. The cohesiveness agent is believed to provide additional hydroxyl groups, which enhance the bonding between grain proteins within the matrix granule. It is further believed that the additional hydroxyl groups so function by enhancing the hydrogen bonding of proteins to starch and to other proteins. The cohesiveness agent may be present in any amount suitable to enhance the cohesiveness of the granules of the enzyme delivery matrix. Suitable cohesiveness agents include one or more of dextrans, maltodextrins, starches, such as corn starch, flours, cellulose, hemicellulose, and the like. For example, the percentage of grain germ and cohesiveness agent in the matrix (not including the enzyme) is 78% corn germ meal and 20% by weight of corn starch.

Because the enzyme-releasing matrix of the invention is made from biodegradable materials, the matrix may be subject to spoilage, such as by molding. To prevent or inhibit such molding, the matrix may include a mold inhibitor, such as a propionate salt, which may be present in any amount sufficient to inhibit the molding of the enzyme-releasing matrix, thus providing a delivery matrix in a stable formulation that does not require refrigeration.

The xylanase enzyme contained in the invention enzyme delivery matrix and methods is preferably a thermostable xylanase, as described herein, so as to resist inactivation of the xylanase during manufacture where elevated temperatures and/or steam may be employed to prepare the palletized enzyme delivery matrix. During digestion of feed containing the invention enzyme delivery matrix, aqueous digestive fluids will cause release of the active enzyme. Other types of thermostable enzymes and nutritional supplements that are thermostable can also be incorporated in the delivery matrix for release under any type of aqueous conditions.

A coating can be applied to the invention enzyme matrix particles for many different purposes, such as to add a flavor or nutrition supplement to animal feed, to delay release of animal feed supplements and enzymes in gastric conditions, and the like. Or, the coating may be applied to achieve a functional goal, for example, whenever it is desirable to slow release of the enzyme from the matrix particles or to control the conditions under which the enzyme will be released. The composition of the coating material can be such that it is selectively broken down by an agent to which it is susceptible (such as heat, acid or base, enzymes or other chemicals). Alternatively, two or more coatings susceptible to different such breakdown agents may be consecutively applied to the matrix particles.

The invention is also directed towards a process for preparing an enzyme-releasing matrix. In accordance with the invention, the process comprises providing discrete plural particles of a grain-based substrate in a particle size suitable for use as an enzyme-releasing matrix, wherein the particles comprise a xylanase enzyme encoded by an amino acid sequence of Group B amino acid sequences or at least 30 consecutive amino acids thereof. Preferably, the process includes compacting or compressing the particles of enzyme-releasing matrix into granules, which most preferably is accomplished by pelletizing. The mold inhibitor and cohesiveness agent, when used, can be added at any suitable time, and preferably are mixed with the grain-based substrate in the desired proportions prior to pelletizing of the grain-based substrate. Moisture content in the pellet mill feed preferably is in the ranges set forth above with respect to the moisture content in the finished product, and preferably is about 14-15%. Preferably, moisture is added to the feedstock in the form of an aqueous preparation of the enzyme to bring the feedstock to this moisture content. The temperature in the pellet mill preferably is brought to about 82°C with steam. The pellet mill may be operated under any conditions that impart sufficient work to the feedstock to provide pellets. The pelleting process itself is a cost-effective process for removing water from the enzyme-containing composition.

In one aspect, the pellet mill is operated with a 1/8 in. by 2 in. die at 100 lb./min. pressure at 82°C. to provide pellets, which then are crumbled in a pellet mill crumbler to provide discrete plural particles having a particle size capable of passing through an 8 mesh screen but being retained on a 20 mesh screen.

The thermostable xylanases of the invention can be used in the pellets of the invention. They can have high optimum temperatures and high heat resistance such that an enzyme reaction at a temperature not hitherto carried out can be achieved. The gene encoding the xylanase according to the present invention (e.g. as set forth in any of the

sequences in Group A nucleic acid sequences) can be used in preparation of xylanases (e.g. using GSSM™ as described herein) having characteristics different from those of the xylanases of Group B amino acid sequences (in terms of optimum pH, optimum temperature, heat resistance, stability to solvents, specific activity, affinity to substrate, secretion ability, translation rate, transcription control and the like). Furthermore, a polynucleotide of Group A nucleic acid sequences may be employed for screening of variant xylanases prepared by the methods described herein to determine those having a desired activity, such as improved or modified thermostability or thermotolerance. For example, U.S. Patent No. 5,830,732, describes a screening assay for determining thermotolerance of a xylanase.

10 *Waste treatment*

The xylanases of the invention can be used in a variety of other industrial applications, e.g., in waste treatment. For example, in one aspect, the invention provides a solid waste digestion process using xylanases of the invention. The methods can comprise reducing the mass and volume of substantially untreated solid waste. Solid waste can be treated with an enzymatic digestive process in the presence of an enzymatic solution (including xylanases of the invention) at a controlled temperature. This results in a reaction without appreciable bacterial fermentation from added microorganisms. The solid waste is converted into a liquefied waste and any residual solid waste. The resulting liquefied waste can be separated from said any residual solidified waste. See e.g., U.S. Patent No. 5,709,796.

20 *Oral care products*

The invention provides oral care product comprising xylanases of the invention. Exemplary oral care products include toothpastes, dental creams, gels or tooth powders, odontics, mouth washes, pre- or post brushing rinse formulations, chewing gums, lozenges, or candy. See, e.g., U.S. Patent No. 6,264,925.

25 *Brewing and fermenting*

The invention provides methods of brewing (e.g., fermenting) beer comprising xylanases of the invention. In one exemplary process, starch-containing raw materials are disintegrated and processed to form a malt. A xylanase of the invention is used at any point in the fermentation process. For example, xylanases of the invention can be used in the processing of barley malt. The major raw material of beer brewing is barley malt. This can be a three stage process. First, the barley grain can be steeped to increase water content, e.g., to around about 40%. Second, the grain can be germinated by incubation at 15 to 25°C for 3 to 6 days when enzyme synthesis is stimulated under the control of gibberellins. In one

aspect, xylanases of the invention are added at this (or any other) stage of the process. Xylanases of the invention can be used in any beer or alcoholic beverage producing process, as described, e.g., in U.S. Patent No. 5,762,991; 5,536,650; 5,405,624; 5,021,246; 4,788,066.

In one aspect, an enzyme of the invention is used to improve filterability and wort viscosity and to obtain a more complete hydrolysis of endosperm components. Use of an enzyme of the invention would also increase extract yield. The process of brewing involves germination of the barley grain (malting) followed by the extraction and the breakdown of the stored carbohydrates to yield simple sugars that are used by yeast for alcoholic fermentation. Efficient breakdown of the carbohydrate reserves present in the barley endosperm and brewing adjuncts requires the activity of several different enzymes.

In one aspect, an enzyme of the invention has activity in slightly acidic pH (e.g., 5.5-6.0) in, e.g., the 40°C to 70°C temperature range; and, in one aspect, with inactivation at 95°C. Activity under such conditions would be optimal, but are not an essential requirement for efficacy. In one aspect, an enzyme of the invention has activity between 40-75° C, and pH 5.5-6.0; stable at 70° for at least 50 minutes, and, in one aspect, is inactivated at 96-100 °C. Enzymes of the invention can be used with other enzymes, e.g., beta-1,4-endoglucanases and amylases.

Medical and research applications

Xylanases of the invention can be used as antimicrobial agents due to their bacteriolytic properties. Xylanases of the invention can be used to eliminating or protecting animals from salmonellae, as described in e.g., PCT Application Nos. WO0049890 and WO9903497.

Other industrial applications

Xylanases of the invention can be used, including Group B amino acid sequences are used in a wide variety of food, animal feed and beverage applications. New xylanases are discovered by screening existing libraries and DNA libraries constructed from diverse mesophilic and moderately thermophilic locations as well as from targeted sources including digestive flora, microorganisms in animal waste, soil bacteria and highly alkaline habitats. Biotrap and primary enrichment strategies using arabinoxylan substrates and/or non-soluble polysaccharide fractions of animal feed material are also useful.

Two screening formats (activity-based and sequence-based) are used in the discovery of novel xylanases. The activity-based approach is direct screening for xylanase activity in agar plates using a substrate such as AZO-xylan (Megazyme). Alternatively a

sequence-based approach may be used, which relies on bioinformatics and molecular biology to design probes for hybridization and biopanning. See, for example, U.S. Patents No. 6,054,267, 6,030,779, 6,368,798, 6,344,328. Hits from the screening are purified, sequenced, characterized (for example, determination of specificity, temperature and pH optima),
5 analyzed using bioinformatics, subcloned and expressed for basic biochemical characterization. These methods may be used in screening for xylanases useful in a myriad of applications, including dough conditioning and as animal feed additive enzymes.

In characterizing enzymes obtained from screening, the exemplary utility in dough processing and baking applications may be assessed. Characterization may include,
10 for example, measurement of substrate specificity (xylan, arabinoxylan, CMC, BBG), temperature and pH stability and specific activity. A commercial enzyme may be used as a benchmark. In one aspect, the enzymes of the invention have significant activity at pH ≥ 7 and 25-35° C, are inactive on insoluble xylan, are stable and active in 50-67% sucrose.

In another aspect, utility as feed additives may be assessed from
15 characterization of candidate enzymes. Characterization may include, for example, measurement of substrate specificity (xylan, arabinoxylan, CMC, BBG), temperature and pH stability, specific activity and gastric stability. In one aspect the feed is designed for a monogastric animal and in another aspect the feed is designed for a ruminant animal. In one aspect, the enzymes of the invention have significant activity at pH 2-4 and 35-40°C, a half-
20 life greater than 30 minutes in gastric fluid, formulation (in buffer or cells) half-life greater than 5 minutes at 85°C and are used as a monogastric animal feed additive. In another aspect, the enzymes of the invention have one or more of the following characteristics: significant activity at pH 6.5-7.0 and 35-40°C, a half-life greater than 30 minutes in rumen fluid, formulation stability as stable as dry powder and are used as a ruminant animal feed
25 additive.

Enzymes are reactive toward a wide range of natural and unnatural substrates, thus enabling the modification of virtually any organic lead compound. Moreover, unlike traditional chemical catalysts, enzymes are highly enantio- and regio-selective. The high degree of functional group specificity exhibited by enzymes enables one to keep track of each
30 reaction in a synthetic sequence leading to a new active compound. Enzymes are also capable of catalyzing many diverse reactions unrelated to their physiological function in nature. For example, peroxidases catalyze the oxidation of phenols by hydrogen peroxide. Peroxidases can also catalyze hydroxylation reactions that are not related to the native function of the

enzyme. Other examples are xylanases which catalyze the breakdown of polypeptides. In organic solution some xylanases can also acylate sugars, a function unrelated to the native function of these enzymes.

The present invention exploits the unique catalytic properties of enzymes.

5 Whereas the use of biocatalysts (i.e., purified or crude enzymes, non-living or living cells) in chemical transformations normally requires the identification of a particular biocatalyst that reacts with a specific starting compound, the present invention uses selected biocatalysts and reaction conditions that are specific for functional groups that are present in many starting compounds. Each biocatalyst is specific for one functional group, or several related
10 functional groups and can react with many starting compounds containing this functional group. The biocatalytic reactions produce a population of derivatives from a single starting compound. These derivatives can be subjected to another round of biocatalytic reactions to produce a second population of derivative compounds. Thousands of variations of the original compound can be produced with each iteration of biocatalytic derivatization.

15 Enzymes react at specific sites of a starting compound without affecting the rest of the molecule, a process which is very difficult to achieve using traditional chemical methods. This high degree of biocatalytic specificity provides the means to identify a single active compound within the library. The library is characterized by the series of biocatalytic reactions used to produce it, a so-called "biosynthetic history". Screening the library for
20 biological activities and tracing the biosynthetic history identifies the specific reaction sequence producing the active compound. The reaction sequence is repeated and the structure of the synthesized compound determined. This mode of identification, unlike other synthesis and screening approaches, does not require immobilization technologies and compounds can be synthesized and tested free in solution using virtually any type of screening assay. It is
25 important to note, that the high degree of specificity of enzyme reactions on functional groups allows for the "tracking" of specific enzymatic reactions that make up the biocatalytically produced library.

Many of the procedural steps are performed using robotic automation enabling the execution of many thousands of biocatalytic reactions and screening assays per day as
30 well as ensuring a high level of accuracy and reproducibility. As a result, a library of derivative compounds can be produced in a matter of weeks which would take years to produce using current chemical methods. (For further teachings on modification of molecules, including small molecules, see PCT/US94/09174).

The invention will be further described with reference to the following examples; however, it is to be understood that the invention is not limited to such examples.

EXAMPLES

5 EXAMPLE 1: PLATE BASED ENDOGLYCOSIDASE ENZYME DISCOVERY: EXPRESSION SCREENING

Titer determination of Lambda Library: Add 1.0 μ L of Lambda Zap Express amplified library stock to 600 μ L *E. coli* MRF' cells ($OD_{600}=1.0$). Dilute MRF' stock with 10mM
10 $MgSO_4$. Incubate mixture at 37°C for 15 minutes, then transfer suspension to 5-6mL of NZY top agar at 50 °C and gently mix. Immediately pour agar solution onto large (150mm) NZY media plate and allow top agar to solidify completely (approximately 30 minutes). Invert the plate. Incubate the plate at 39°C for 8-12 hours. (The number of plaques is approximated. Phage titer determined to give 50,000 pfu/plate. Dilute an aliquot of Library phage with SM
15 buffer if needed.)

Substrate screening: Add Lambda Zap Express (50,000 pfu) from amplified library to 600 μ L of *E. coli* MRF' cells ($OD_{600}=1.0$) and incubate at 37°C for 15 minutes. While phage/cell suspension is incubating, add 1.0mL of desired polysaccharide dye-labeled substrate (usually 1-2% w/v) to 5.0mL NZY top agar at 50°C and mix thoroughly. (Solution kept at 50°C until
20 needed.) Transfer the cell suspension to substrate/top agar solution and gently mix. Immediately pour solution onto large (150mm) NZY media plate. Allow top agar to solidify completely (approximately 30 minutes), then invert plate. Incubate plate at 39°C for 8-12 hours. Observe plate for clearing zones (halos) around plaques. Core plaques with halos out of agar and transfer to a sterile micro tube. (A large bore 200 μ L pipette tip works well to
25 remove (core) the agar plug containing the desired plaque.) Resuspend phage in 500 μ L SM buffer. Add 20 μ L chloroform to inhibit any further cell growth.

Isolation of pure clones: Add 5 μ L of resuspended phage suspension to 500 μ L of *E. coli* MRF' cells ($OD_{600}=1.0$). Incubate at 37°C for 15 minutes. While phage/cell suspension is incubating, add 600 μ L of desired polysaccharide dye-labeled substrate (usually 1-2% w/v) to
30 3.0mL NZY top agar at 50°C and mix thoroughly. (Solution kept at 50°C until needed.) Transfer cell suspension to substrate/top agar solution and gently mix. Immediately pour solution onto small (90mm) NZY media plate and allow top agar to solidify completely

(approximately 30 minutes), then invert plate. Incubate plate at 39°C for 8-12 hours. Plate observed for a clearing zone (halo) around a single plaque (pure clone). (If a single plaque cannot be isolated, adjust titer and replate phage suspension.) Phage are resuspended in 500µL SM buffer and 20µL Chloroform is added to inhibit any further cell growth.

- 5 Excision of pure clone: Allow pure phage suspension to incubate at room temperature for 2 to 3 hours or overnight at 4°C. Add 100µL of pure phage suspension to 200µL *E. coli* MRF' cells (OD₆₀₀=1.0). Add 1.0µL of ExAssist helper phage (>1 x 10⁶ pfu/mL; Stratagene). Incubate suspension at 37°C for 15 minutes. Add 3.0 mL of 2 x YT media to cell suspension. Incubate at 37°C for 2-2.5 hours while shaking. Transfer tube to 70°C for 20 minutes.
- 10 Transfer 50-100 µL of phagemid suspension to a micro tube containing 200µL of *E. coli* Exp 505 cells (OD₆₀₀=1.0). Incubate suspension at 37°C for 45 minutes. Plate 100 µL of cell suspension on LB_{kan 50} media (LB media with Kanamycin 50µg/mL). Incubate plate at 37°C for 8-12 hours. Observe plate for colonies. Any colonies that grow contain the pure phagemid. Pick a colony and grow a small (3-10mL) liquid culture for 8-12 hours. Culture
- 15 media is liquid LB_{kan 50}.

- Activity verification: Transfer 1.0mL of liquid culture to a sterile micro tube. Centrifuge at 13200 rpm (16000 g's) for 1 minute. Discard supernatant and add 200µL of phosphate buffer pH 6.2. Sonicate for 5 to 10 seconds on ice using a micro tip. Add 200 µL of appropriate substrate, mix gently and incubate at 37 °C for 1.5-2 hours. A negative control should also be
- 20 run that contains only buffer and substrate. Add 1.0mL absolute ethanol (200 proof) to suspension and mixed. Centrifuge at 13200 rpm for 10 minutes. Observe supernatant for color. Amount of coloration may vary, but any tubes with more coloration than control is considered positive for activity. A spectrophotometer can be used for this step if so desired or needed. (For Azo-xylan, Megazyme, read at 590nm).

- 25 RFLP of pure clones from same Libraries: Transfer 1.0mL of liquid culture to a sterile micro tube. Centrifuge at 13200 rpm (16000 g's) for 1 minute. Follow QIAprep spin mini kit (Qiagen) protocol for plasmid isolation and use 40 µL holy water as the elution buffer. Transfer 10 µL plasmid DNA to a sterile micro tube. Add 1.5µL Buffer 3 (New England Biolabs), 1.5µL 100X BSA solution (New England Biolabs) and 2.0µL holy water. To this
- 30 add 1.0µL Not 1 and 1.0µL Pst 1 restriction endonucleases (New England Biolabs). Incubate for 1.5 hours at 37°C. Add 3.0µL 6X Loading buffer (Invitrogen). Run 15µL of

digested sample on a 1.0% agarose gel for 1-1.5 hours at 120 volts. View the gel with a gel imager. Perform sequence analysis on all clones with a different digest pattern.

Table 6 describes various properties of exemplary enzymes of the invention.

Table 6

5

SEQ ID NO.	Topt*	Tstab**	pHopt*	Significant activities	pl	M _w	Notes
151, 152	50°C	<1 min at 65°C	5.5-9.0	AZO-xylan	5.7	40.2	
155, 156	50°C	<1 min at 65°C	5.5-8.0	AZO-xylan	8.8	62.7	
169, 170	50°C	> 1 min at 65°C; < 1 min at 85°C	7.0	AZO-xylan	8.7	36.7	
195, 196	50°C	>1 min at 65°C < 10 min, < 1 min 85°C	5.5	AZO-xylan	8.5	36.7	
215, 216	85°C	<3 min at 85°C	5.5-8.0	AZO-xylan	8.6	34.8	
47, 48	50°C	< 0.5 min at 65°C; < 1 min at 85°C	7.0-8.0	AZO-xylan	6.2	40.3	
191, 192	385°C	> 30 sec at 85°C	5.5	AZO-xylan	7.8	34.6	
247, 248	50°C	< 1 min at 65°C	8.0	AZO-xylan	9.4	43.5	
7, 8	50°C	> 1 min 85°C < 5 min	5.5	AZO-xylan	4.5	55.3	
221, 222	50-65°C	<1 min at 75°C	5.5	AZO-xylan	8.3	34.6	
163, 164	65°C	<1 min at 65°C	7.0	AZO-xylan	6.3	36.0	
19, 20	37°C	<5 min at 50°C	7.0 - 8.0	AZO-xylan	9.2	41.5	
87, 88	37 - 50°C	< 1 min at 85°C	8.0	AZO-xylan	5.2	36.7	
81, 82	50°C	< 1 min at 65°C	7.0 - 9.0	AZO-xylan	5.3	38.8	
91, 92	50°C	< 1 min at 65°C	7 - 8	AZO-xylan, AZO-CMC	5.4	39.0	
61, 62	37°C	<5 min at 50°C	7.0 - 9.0	AZO-xylan, AZO-CMC	5.4	40	
159, 160	85°C	< 30 sec at 85°C	5.5	AZO-xylan	8.3	34.5	
233, 234	50°C	> 30 sec < 1 min at 65°C; < 1 min at 85°C	7.0	AZO-xylan	8.5	35.1	
203, 204	50 - 65°C	> 1 min at 65°C < 5 min, < 1 min 85°C	5.5	AZO-xylan	9.5	21.7	
181, 182	385°C	> 1 min at 85°C	5.5-8.0	AZO-xylan	8.8	35.5	
227, 228	65°C	>1 min at 85°C < 5 min	5.5 - 7.0	AZO-xylan	7.8	25.8	
45, 46	345°C	35 min 45°C, <0.5 min 55°C	> 5.5	AZO-xylan	6.7	40.4	***
231, 232	65°C	>10 min at 50°C	5.5 - 7.0	AZO-xylan	8.4	31.4	
129, 130	65°C	<1min at 75°C	5.5	AZO-xylan	5.1	116	
93, 94	50°C	< 1 min at 60°C	8.0 - 9.0	AZO-xylan	5.3	39.1	
189, 190	65°C	<1 min at 65°C	5.5	AZO-xylan	9.2	20.3	****
49, 50	70°C	<20 min 70°C	>5	AZO-xylan	5.7	38.9	
85, 86	50°C	>5 min at 85°C	5.5 - 7.0	AZO-xylan	6.1	48.4	
99, 100	50°C	<1 min at 75°C	5.5 - 8.0	AZO-xylan	10.8	36.6	
123, 124	385°C	<30 sec 100 °C	5.5-7.0	AZO-xylan	6.1	44.1	
249, 250	45°C	>1 min 75°C < 10 min	5.5	AZO-xylan	5.3	93	
167, 168	85°C	< 5 min 85°C	5.5	AZO-xylan	9.5	21.7	
207, 208	75°C	< 5 min 65 °C	5.5	AZO-xylan	9.1	20.4	
251, 252	65-75°C	< 1 min 85 °C	5.5	AZO-xylan	8.8	20.4	*****
11, 12	<90°C	<40 min 70°C	>6	AZO-xylan	6.8	43.9	
177, 178	65°C	< 1 min at 75°C	5.5	AZO-xylan	8.7	44.6	
9, 10	50°C	<1min at 65°C	5.5 - 7.0	AZO-xylan	4.9	46.1	
43, 44	37°C	unstable	5.5-7.0	AZO-xylan	4.9	39.1	
113, 114	65 - 75°C	< 1 min at 75°C	5.5 - 8.0	AZO-xylan	5	41.2	

SEQ ID NO.	Topt*	Tstab**	pHopt*	Significant activities	pI	M _w	Notes
75, 76	50°C	< 1 min 85°C	7.0 - 9.0	AZO-xylan	4.7	39.4	
111, 112	37°C	>10 min 50°C	7 - 8	AZO-xylan	5.6	41.0	
117, 118	37°C	unstable	7-8	AZO-xylan	9.1	53.3	
115, 116	-	-	-	AZO-xylan	8.9	50.8	
125, 126	37°C	-	8.0	AZO-xylan	5.3	41.1	
137, 138	50°C	< 30 sec at 65°C	5.5	AZO-xylan	5.7	38.5	
69, 70	85°C	< 5 min at 85°C	5.5-9.0	AZO-xylan	6.4	58.0	
205, 206	50°C	<1min at 65°C	5.5 - 8	AZO-xylan	4.3	35.1	
211, 212	50°C	<1min at 65°C	5.5	AZO-xylan	4.4	35.4	
197, 198	65°C	<1 min at 65°C	5.5	AZO-xylan	8.8	20.1	
31, 32	37°C	unstable	7.0	AZO-xylan	5.1	54.4	
13, 14	50°C	<1 min at 65°C	7	AZO-xylan	5.5	40.0	
65, 66	50°C	< 1 min at 65°C	5.5	AZO-xylan, AZO-CMC	4.8	55.5	
257, 258	37°C	unstable	5.5	AZO-xylan, AZO-barley β-glucan, AZO-CMC	5.3	100.8	
57, 58	50°C	<1min at 65°C	7.0	AZO-xylan	4.8	56.7	
185, 186	50-75°C	< 1 min at 80°C	5.5	AZO-xylan	8.8	23.2	
243, 244	75°C	>0.5 min @ 85°C	5.5	AZO-xylan	8.8	44.4	
77, 78	50°C	< 5 min at 65°C, < 1 min 85°C	5.5	AZO-xylan	5.3	44.5	
229, 230	37°C	30 min 55°C, < 5 min 75°C	5.5	AZO-xylan	8.7	20.6	*****
109, 110	65°C	>0.5 min @ 75°C	5.5	AZO-xylan	4.9	45.2	
193, 194	65°C	< 1 min at 75°C	5.5	AZO-xylan	5.4	29.1	
173, 174	65°C	< 1 min at 80°C	7.0	AZO-xylan	7.6	51.6	
59, 60	37°C	<1min at 65°C	7.0	AZO-xylan	6.6	42.5	
101, 102	50°C	>0.5 min @ 65°C	7.0	AZO-xylan	8.7	41.1	
55, 56	37°C	> 5 min at 50°C; < 1 min at 85°C	7.0	AZO-xylan	6.5	41.8	
15, 16	50°C	< 1 min at 65°C	7.0	AZO-xylan	6.4	40.2	
131, 132	-	-	-	AZO-xylan	5.6	42.1	
145, 146	65-85°C	< 1 min at 85°C	5.5	AZO-xylan	5.2	43.7	
219, 220	-	-	5.5	AZO-xylan	6.6	34.5	
253, 254	65°C	> .5 min at 85°C	5.5 - 7	AZO-xylan	7.8	34.6	
255, 256	65°C	> 1 min 65°C <3 min	5.5-7.0	AZO-xylan	8.3	35.0	

* pH or temperature optima determined by initial rates using AZO-AZO-xylan as a substrate

** thermal stability, time that enzyme retained significant activity (approx. > 50 %)

*** Dough conditioning

**** GSSM™ parent for thermal tolerance evolution for animal feed applications

***** N35D mutation made to increase low pH activity- based on public knowledge- mutant enzyme's relative activity at pH 4 significantly increased

5 ***** Dough conditioning

EXAMPLE 2: GSSM™ SCREEN FOR THERMAL TOLERANT MUTANTS

- The following example describes an exemplary method for screening for thermally tolerant enzymes.
- 10 Master Plates: Prepare plates for a colony picker by labeling 96 well plates and aliquoting 200 μL LB Amp100 into each well. (~20ml needed per 96 well plate). After the plates are

returned from the picker, remove media from row 6 from plate A. Replace with an inoculation of SEQ ID NO: 189. Place in a humidified 37°C incubator overnight.

Assay Plates: Pin tool cultures into a fresh 96 well plate (200 µL /well LB Amp100).

Remove plastic cover and replace with Gas Permeable Seal. Place in a humidified incubator overnight. Remove the seal and replace plastic lid. Spin cultures down in tabletop centrifuge at 3000 rpm for 10 min. Remove supernatant by inversion onto a paper towel. Aliquot 45 µL Cit-Phos-KCl buffer pH 6 into each well. Replace the plastic lid with an aluminum plate seal. Use a roller to get a good seal. Resuspend cells in a plate shaker at level 6-7 for ~30 seconds.

10 Place the 96 well plate in 80°C incubator for 20 minutes. Do not stack. Thereafter, immediately remove plates to ice water to cool for a few minutes. Remove the aluminum seal and replace with a plastic lid. Add 30 µL of 2 % Azo-xylan. Mix as before on the plate shaker. Incubate 37°C in a humidified incubator overnight.

Add 200 µL ethanol to each well and pipette up and down a couple of times to mix. As an alternative to changing tips each time, rinse in an ethanol wash and dry by expelling into a paper towel. Spin the plates at 3000 rpm for 10 minutes. Remove 100 µL of supernatant to a fresh 96 well plate. Read the OD₅₉₀.

EXAMPLE 3: GSSM™ ASSAY FOR HIT VERIFICATION OF THERMAL TOLERANT MUTANTS

20 The following example describes an exemplary method for assaying for thermally tolerant enzymes.

Pin tool or pick clones into duplicate 96 well plates (200ul /well LB Amp100). Remove the plastic cover and replace with a Gas Permeable Seal. Place in a humidified incubator overnight. Remove the Seal and replace with a plastic lid. Pintool the clones to solid agar. Spin cultures down in tabletop centrifuge at 3000 rpm for 10 min. Remove the supernatant by inversion onto a paper towel. Aliquot 25 µl BPER/Lysozyme/DNase solution (see below) into each well. Resuspend cells in a plate shaker on level 6-7 for ~30 seconds.

Incubate the plate on ice for 15 minutes. Add 20 µL of Cit-Phos-KCl buffer pH 6 into each well. Replace the plastic lid with an aluminum plate seal. Use a roller to get a good seal. Mix on a plate shaker at level 6-7 for ~30 seconds.

Place one 96 well plate in an 80°C incubator for 20 minutes and the other at 37°C. Do not stack. Immediately remove the plates to watery ice to cool for a few minutes (use a large plastic tray if needed). Remove the aluminum seal. Add 30 µl of 2% Azo-xylan.

Seal with a plastic gas permeable seal. Mix as before on the plate shaker. Incubate a set of 37°C and 80°C plates in humidified incubator at 37°C for 2 hours and another set for 4 hours.

After incubation, let the plate sit for ~5 minutes at room temperature. Add 200 µL ethanol to each well and pipette up and down a couple of times to mix. Instead of changing tips each time, rinse in an ethanol wash and dry by expelling into a paper towel. But, use a new set of tips for each clone. Spin plates at 3000 rpm 10 minutes. Remove 100 µL of supernatant to a fresh 96 well plate. Read OD₅₉₀.

BPER/Lysozyme/DNase solution (4.74 mL total):

4.5 mL BPR

200 µL 10 mg/mL Lysozyme (made fresh in pH 6 Cit-phos-buffer)

40 µL 5 mg/mL DNase I (made fresh in pH 6 Cit-phos buffer)

EXAMPLE 4: Xylanase assay with wheat arabinoxylan as substrate

The following example describes an exemplary xylanase assay that can be used, for example, to determine is an enzyme is within the scope of the invention.

SEQ ID NOS: 11, 12, 69, 70, 77, 78, 113, 114, 149, 150, 159, 160, 163, 164, 167, 168, 181, 182, 197, and 198 were subjected to an assay at pH 8 (Na-phosphate buffer) and 70°C using wheat arabinoxylan as a substrate. The enzymes were characterized as set forth in Table 7.

Table 7

SEQ ID NOS:	Protein Concentration (mg/ml)	volume of lysate added to each vial	#of vials	Units/ml*	protein (mg/mL)	U/mg
11, 12	42	0.5	10	163	22.0	7.4
113, 114	37	0.6	10	66	22.0	3.0
163, 164	35	0.6	10	25	22.0	1.1
197, 198	23	1.0	10	31	22.0	1.4
167, 168	10	2.2	10	228	22.0	10.4
77, 78	47	0.5	10	29	22.0	1.3
69, 70	18	1.3	10	36	22.0	1.7
181, 182	28	0.8	10	24	22.0	1.1
159, 160	25	0.9	10	43	22.0	2.0
149, 150	42	0.5	10	24	22.0	1.1

*Based on addition of 1 mL of water to each sample.

Units are umoles xylose released per minute based on a reducing sugar assay.

EXAMPLE 5: Generation of an exemplary xylanase of the invention

The following example describes the generation of an exemplary xylanase of the invention using gene site-saturation mutagenesis (GSSM™) technology, designated the

“9x” variant or mutant (the nucleic acid as set forth in SEQ ID NO:377, the polypeptide sequence as set forth in SEQ ID NO:378).

GSSM™ was used to create a comprehensive library of point mutations in the exemplary SEQ ID NO:190, “wild-type” xylanase (encoded by SEQ ID NO:189). The xylanase thermotolerance screen described above identified nine single site amino acid mutants (Figure 6A) (D8F, Q11H, N12L, G17I, G60H, P64V, S65V, G68A & S79P) that had improved thermal tolerance relative to the wild type enzyme (as measured following a heat challenge at 80°C for 20 minutes). Wild-type enzyme and all nine single site amino acid mutants were produced in *E. coli* and purified utilizing an N-terminal hexahistidine tag.

There was no noticeable difference in activity due to the tag.

Figure 6 illustrates the nine single site amino acid mutants of “variant 9x”, or, as set forth in SEQ ID NO:378 (encoded by SEQ ID NO:377), as generated by Gene Site Saturation Mutagenesis (GSSM™) of the exemplary SEQ ID NO:190 “wild-type” enzyme (encoded by SEQ ID NO:189). Figure 6A is a schematic diagram illustrating position, numbering and the amino acid change for the thermal tolerant point mutants of the “wild-type” gene (SEQ ID NO:190, encoded by SEQ ID NO:189). A library of all 64 codons was generated for every amino acid position in the gene (~13,000 mutants) and screened for mutations that increased thermal tolerance. The “9X” variant was generated by combining all 9 single-site mutants into one enzyme. The corresponding melting temperature transition midpoint (T_m) determined by DSC for each mutant enzyme and the “9X” (SEQ ID NO:378) variant is shown on the right. Figure 6B illustrates the unfolding of the “wild-type” (SEQ ID NO:190) and “9X” (SEQ ID NO:378) “variant/mutant” enzymes was monitored by DSC at a scan rate of 1°C/min. Baseline subtracted DSC data were normalized for protein concentration.

Xylanase activity assays

Enzymatic activities were determined using 400 μ L of 2% Azo-xylan as substrate in 550 μ L of CP (citrate-phosphate) buffer, pH 6.0 at the indicated temperatures. Activity measurements as a function of pH were determined using 50 mM Britton and Robinson buffer solutions (pH 3.0, 5.0, 6.0, 7.0, 8.0 and 9.0) prepared by mixing solutions of 0.1 M phosphoric acid solution, 0.1 M boric acid and 0.1 M acetic acid followed by pH adjustment with 1 M sodium hydroxide. Reactions were initiated by adding 50 μ L of 0.1 mg/ml of purified enzyme. Time points were taken from 0 to 15 minutes where 50 μ L of reaction mixture was added to 200 μ L of precipitation solution (100% ethanol). When all

time points had been taken, samples were mixed, incubated for 10 minutes and centrifuged at 3000 g for 10 minutes at 4°C. Supernatant (150 μ L) was aliquoted into a fresh 96 well plate and absorbance was measured at 590 nm. A₅₉₀ values were plotted against time and the initial rate was determined from the slope of the line.

5 *Differential Scanning Calorimetry (DSC).*

Calorimetry was performed using a Model 6100 Nano II DSC apparatus (Calorimetry Sciences Corporation, American Fork, UT) using the DSCRun software package for data acquisition, CpCalc for analysis, CpConvert for conversion into molar heat capacity from microwatts and CpDeconvolute for deconvolution. Analysis was carried out with 1mg/ml recombinant protein in 20 mM potassium phosphate (pH 7.0) and 100 mM KCl at a scan rate of 1°C/min. A constant pressure of 5 atm was maintained during all DSC experiments to prevent possible degassing of the solution on heating. The instrumental baseline was recorded routinely before the experiments with both cells filled with buffer. Reversibility of the thermally induced transitions was tested by reheating the solution in the calorimeter cell immediately after cooling the first run.

Thermal tolerance determination.

All enzymes were analyzed for thermal tolerance at 80°C in 20 mM potassium phosphate (pH 7.0) and 100 mM KCl. The enzymes were heated at 80°C for 0, 5, 10 or 30 minutes in thin-walled tubes and were cooled on ice. Residual activities were determined with Azo-xylan as substrate using the assay described above for activity measurement.

Polysaccharide Fingerprinting.

Polysaccharide fingerprints were determined by polysaccharide analysis using carbohydrate gel electrophoresis (PACE). Beechwood xylan (0.1 mg/mL, 100 μ L, Sigma, Poole, Dorset, UK) or xylooligosaccharides (1 mM, 20 μ L, Megazyme, Wicklow, Ireland) were treated with enzyme (1 – 3 μ g) in a total volume of 250 μ L for 16 hours. The reaction was buffered in 0.1 M ammonium acetate pH 5.5. Controls without substrates or enzymes were performed under the same conditions to identify any unspecific compounds in the enzymes, polysaccharides/oligosaccharides or labeling reagents. The reactions were stopped by boiling for 20 min. Assays were independently performed at least 2 times for each condition. Derivatization using ANTS (8-aminonaphthalene-1,3,6-trisulfonic acid, Molecular Probes, Leiden, The Netherlands), electrophoresis and imaging were carried out as described (Goubet, F., Jackson, P., Deery, M. and Dupree, P. (2002) *Anal. Biochem.* 300, 53–68).

Fitness Calculation.

The fitness (F_n), for a given enzyme variant, n , was calculated by equally weighting increase in denaturation temperature transition midpoint (T_m) and increase (or decrease) in enzymatic activity relative to the largest difference in each parameter across all variants: $F_n = F_{Tn} + F_{Vn}$, where $F_{Tn} = T_m$ fitness factor of the variant and F_{Vn} = activity fitness factor of the variant. The fitness factors for each (T_m and activity) are relative to the largest difference in T_m or rate across all of the variants. $F_{Tn} = (T_m - T_{mL}) / (T_{mH} - T_{mL})$ where T_{mn} is the T_m for the given variant, n , and T_{mL} is the lowest T_m across all variants and T_{mH} the highest T_m across all variants and $F_{Vn} = (V_n - V_L) / (V_H - V_L)$ where V_n is the relative rate for the given variant, n , and V_L is the lowest rate across all variants and V_H the highest rate across all variants.

Evolution by the GSSM™ method.

GSSM™ technology was used to create a comprehensive library of point mutations in the exemplary xylanase of the invention SEQ ID NO:190 (encoded by SEQ ID NO:189); including the exemplary xylanase of the invention SEQ ID NO:378 (encoded by SEQ ID NO:377). The xylanase thermotolerance screen described above identified nine single site amino acid mutants (Figure 6A), D8F, Q11H, N12L, G17I, G60H, P64V, S65V, G68A & S79P, that had improved thermal tolerance relative to the exemplary “wild type” enzyme SEQ ID NO:190 (encoded by SEQ ID NO:189), as measured following a heat challenge at 80°C for 20 minutes. Wild-type enzyme and all nine single site amino acid mutants were produced in *E. coli* and purified utilizing an N-terminal hexahistidine tag. There was no noticeable difference in activity due to the tag.

To determine the effect of the single amino acid mutations on enzymatic activity, all nine mutants were purified and their xylanase activity (initial rates at the wild-type temperature optimum, 70°C) was compared to that of the exemplary SEQ ID NO:190 “wild-type” enzyme. Enzyme activities were comparable to wild type (initial rate normalized to 1.0) for D8F, N12L, G17I, G60H, P64V, S65V G68A and S79P mutants (relative initial rates 0.65, 0.68, 0.76, 1.1, 1.0, 1.2, 0.98 and 0.84 respectively) confirming that these mutations do not significantly alter the enzymatic activity. Initial rates were measured 3 or more times and variance was typically less than 10 %. In contrast to these eight mutants, a notable reduction in enzymatic activity was observed for the best thermal tolerant, single site mutant, Q11H (relative initial rate 0.35).

Melting temperature (T_m) of "wild-type" and thermal tolerant single site amino acid mutant enzymes.

The purified SEQ ID NO:190 "wild-type" xylanase and the nine thermal tolerant single site amino acid mutants were analyzed using differential scanning calorimetry (DSC). Aggregation was apparent for the wild-type enzyme as evidenced by a shoulder in the DSC trace for its thermal denaturation, see Figure 6B. The evolved mutant enzymes showed no indication of aggregation. For all enzymes, thermally induced denaturation was irreversible and no discernible transition was observed in a second scan of the sample. Due to the irreversibility of denaturation, only the apparent T_m (melting temperature) could be calculated (as described, e.g., by Sanchez-Ruiz (1992) *Biophys. J.* 61:921–935; Beldarrain (2000) *Biotechnol. Appl. Biochem.* 31:77–84). The T_m of the wild-type enzyme was 61°C while the T_m 's of all point mutants were increased and ranged from 64°C to 70°C (Figure 6A). The Q11H mutation introduced the largest increase (T_m = 70°C) over wild-type followed by P64V (69°C), G17I (67°C) and D8F (67°C).

15 *The "9X" combined GSSM™ exemplary enzyme SEQ ID NO:378*

The "9X" enzyme (SEQ ID NO:378) was constructed by combining the single-site changes of the nine thermal tolerant up-mutants by site-directed mutagenesis (Figure 6A). The "9X" (SEQ ID NO:378) enzyme was expressed in *E. coli* and purified to homogeneity. DSC was performed to determine the melting temperature. The T_m of "9X" enzyme was 34 degrees higher than SEQ ID NO:190, the "wild-type" enzyme, demonstrating a dramatic shift in its thermal stability (Figure 6B).

To evaluate the effect of the combined mutations and elevated melting temperature on the enzyme's biochemical properties, pH and temperature profiles were constructed and compared to SEQ ID NO:190, the "wild-type" enzyme. Figure 7 illustrates the biochemical characterization of "wild type" and "evolved" 9X mutant enzymes. Figure 7A illustrates the pH-dependence of activity for the wild-type and evolved 9X mutant enzymes. Xylanase activity was measured at 37°C at each pH and the initial velocity was plotted against absorbance at 590 nm to determine initial rates. Figure 7B illustrates the temperature-dependence of activity for the wild-type and evolved 9X mutant enzymes. The optimum temperatures of the wild-type and 9X mutant enzymes were measured over a temperature range of 25-100°C at pH 6.0 and are based on initial rates measured over 5 minutes. Figure 7C illustrates the thermal stability of wild-type and evolved 9X mutant enzymes. Thermal dependence of activity of the wild-type and evolved 9X mutant enzymes was measured by first heating the enzymes at each of the indicated temperatures for 5

minutes followed by cooling to room temperature and the measurement of residual activity (initial rate at 37°C, pH 6.0). For all experiments initial rates were measured 2 or more times and the variation was less than 10 %.

SEQ ID NO:190 and SEQ ID NO:378 (the "9X" mutant) enzyme had
5 comparable pH/activity profiles with the highest activity between pH 5 and 6 (Figure 7A). Both enzymes had similar initial rate/temperature optima at 70°C, however, SEQ ID NO:190, the "wild-type" enzyme had higher activity at lower temperatures (25-50°C) whereas SEQ ID NO:378 (the "9X" mutant) retained more than 60% of its activity up to 100°C (determined by initial rate) in the presence of substrate (Figure 7B). The activity of SEQ ID NO:190, the
10 "wild-type" enzyme was not detectable at temperatures above 70°C.

To determine the effect of the 9 combined mutations on enzyme thermal tolerance, residual activity was measured and compared to SEQ ID NO:190, the "wild-type" enzyme. Residual activity was determined by a heat challenge for 5 minutes at each temperature (37, 50, 60, 70, 80 and 90°C) followed by activity measurements at 37°C. SEQ
15 ID NO:190 was completely inactivated above 70°C while the evolved 9X mutant displayed significant activity after heating at 70, 80 and even 90°C (Figure 7C). Furthermore, although the activity of the wild-type enzyme decreased with increasing temperature, the 9X variant was somewhat activated by heating at temperatures up to 60°C.

Generation of combinatorial GSSM™ variants using GeneReassembly™ technology.

20 To identify combinatorial variants of the 9 single site amino acid mutants with highest thermal tolerance and activity compared to the additively constructed SEQ ID NO:378 (the "9X" variant), a GeneReassembly™ library (U.S. Patent No. 6,537,776) of all possible mutant combinations (2^9) was constructed and screened. Using thermal tolerance as the screening criterion, 33 unique combinations of the nine mutations were identified as was
25 the original 9X variant. A secondary screen was performed to select for variants with higher activity/expression than the evolved 9X. This screen yielded 10 variants with sequences possessing between 6 and 8 of the original single mutations in various combinations, as illustrated in Figure 8A. Figure 8 illustrates the combinatorial variants identified using GeneReassembly™ technology. Figure 8A illustrates the GeneReassembly™ library of all
30 possible combinations of the 9 GSSM™ point mutations that was constructed and screened for variants with improved thermal tolerance and activity. Eleven variants including the 9X variant were obtained. As shown in the figure, the variants possessed 6, 7, 8, or 9 of the point mutations in various combinations. The corresponding melting temperature transition

midpoint (T_m) determined by DSC of each variant is shown on the right. Figure 8B illustrates the relative activity (initial rate measured over a 5 minute time period) of the 6X-2 and 9X variants compared to wild-type at the temperature optimum (70°C) and pH 6.0. Error bars show the range in the initial rate for 3 measurements.

5 The melting temperature (T_m) of each of the combinatorial variants was at least 28°C higher than wild type (Figure 8A) and all of the reassembly variants displayed higher relative activity than the 9X enzyme. The activity of one variant in particular, 6X-2, was greater than the wild-type enzyme and significantly better (1.7X) than the 9X enzyme (Figure 8B). Sequence comparison of the reassembly variants identified at least 6 mutations
10 that were required for the enhanced thermostability (>20 degrees). All 33 unique variants found in the initial thermostability screen contained both Q11H and G17I mutations demonstrating their importance for thermal tolerance.

Analysis of wild-type and variant polysaccharide product fingerprints.

 The products generated by the “wild-type,” 6X-2 and 9X variants were
15 compared by polysaccharide analysis using carbohydrate gel electrophoresis (PACE). Different substrates (oligosaccharides and polysaccharides) were tested for hydrolysis by the xylanases. The digestion products of the 3 xylanases tested were very similar, as illustrated in Figure 9. All three enzymes hydrolyzed $(\text{Xyl})_6$ and $(\text{Xyl})_5$, mainly into both $(\text{Xyl})_3$ and $(\text{Xyl})_2$, and $(\text{Xyl})_4$ was hydrolyzed to $(\text{Xyl})_2$ (Figure 9A). Only a small amount of hydrolysis
20 of $(\text{Xyl})_3$ into $(\text{Xyl})_2$ and Xyl was observed indicating that $(\text{Xyl})_3$ is a relatively poor substrate for the enzyme. No activity was detected on $(\text{Xyl})_2$. Beechwood xylan, which contains glucuronosyl residues, was hydrolyzed by all three enzymes mainly into $(\text{Xyl})_2$ and $(\text{Xyl})_3$, but other bands were detected that migrated between oligoxylan bands (Figure 9B). In PACE analysis, each oligosaccharide has a specific migration depending on the sugar composition
25 and degree of polymerization (Goubet, F., Jackson, P., Deery, M. and Dupree, P. (2002) *Anal. Biochem.* 300, 53–68), thus, these bands likely correspond to oligoglucuronoxylans. Therefore, the evolved enzymes retained the substrate specificity of the “wild-type” enzyme.

 As noted above, Figure 9 illustrates the product fingerprints of “wild-type” SEQ ID NO:190 (encoded by SEQ ID NO:189), 6X-2 (SEQ ID NO:380, encoded by SEQ ID
30 NO:379) and SEQ ID NO:378 (the “9X” mutant) enzyme variant, as determined by PACE. Figure 9A illustrates fingerprints obtained after hydrolysis of oligoxylans $(\text{Xyl})_3$, $(\text{Xyl})_4$, $(\text{Xyl})_5$ and $(\text{Xyl})_6$ by “wild-type” and variant enzymes. Control lanes contain oligosaccharide incubated under the assay conditions in the absence of enzyme. Figure 9B illustrates the

fingerprints obtained after hydrolysis of Beechwood xylan by wild-type and variant enzymes. Standards contained (Xyl)₂, (Xyl)₃, (Xyl)₄. All assays were performed at 37°C and pH 5.5.

A combination of laboratory gene evolution strategies was used to rapidly generate a highly active, thermostable xylanase optimized for process compatibility in a number of industrial market applications. GSSM™ methodology was employed to scan the entire sequence of the exemplary “wild type” xylanase SEQ ID NO:190 (encoded by SEQ ID NO:189) and to identify 9 point mutations that improve its thermal tolerance. Although it had no discernable effect on the hydrolysis product profile of the enzyme, as illustrated in Figure 9, the addition of the 9 mutations to the protein sequence resulted in a moderate reduction in enzymatic specific activity at SEQ ID NO:190 (the “wild-type”)’s temperature optimum, 70°C, see Figure 9B. Using the GeneReassembly™ method to generate a combinatorial library of the 9 single site amino acid mutants, this reduction in activity was overcome. Ten thermostable variants (T_m’s between 89°C and 94°C) with activity better than the “9X” variant were obtained from screening the GeneReassembly™ library. With a T_m of 90°C, enzymatic specific activity surpassing wild-type and a product fingerprint unaltered and comparable to SEQ ID NO:190 (the “wild-type”), the 6X-2 variant (SEQ ID NO:380, encoded by SEQ ID NO:379) is particularly notable. To our knowledge the shift in T_m obtained for these variants is the highest increase reported from the application of directed evolution technologies.

SEQ ID NO:380 (the 6X-2 variant) includes the following changes, as compared to SEQ ID NO:190 (the “wild-type”): D8F, Q11H, G17I, G60H, S65V and G68A. SEQ ID NO:379 includes the following nucleotide changes, as compared to the “wild type” SEQ ID NO:189: the nucleotides at positions 22 to 24 are TTC, the nucleotides at positions 31 to 33 are CAC, the nucleotides at positions 49 to 51 are ATA, the nucleotides at positions 178 to 180 are CAC, the nucleotides at positions 193 to 195 are GTG, the nucleotides at positions 202 to 204 are GCT.

In order to gauge the effectiveness of combinatorial mixing versus addition of the point mutants to the desired phenotype, a fitness parameter combining contributions both from changes in enzyme activity and thermostability was calculated for each mutant. The term fitness as described here is not an objective measure that can be compared to other enzymes, but rather a term that allows the measurement of the success of directed evolution of this particular xylanase. Since enzyme fitness, F, is calculated by equally weighting changes in T_m and enzyme activity for this set of variants, the maximum allowable fitness

value is 2 ($F_T \leq 1$ and $F_V \leq 1$, see above). In other words, if the variant with the best activity also had the highest T_m , its fitness value would be 2. With a fitness value near 2 (see Fig. 10B), the 6X-2 variant (SEQ ID NO:380, encoded by SEQ ID NO:379) is the closest to possessing the best possible combination of thermal stability and enzyme activity. The single
5 site mutation that confers the highest value of fitness is S65V. Although the T_m of the S65V mutant is lower than that of the Q11H mutant (66°C versus 70°C respectively), it has a higher fitness value since its specific activity is not reduced relative to wild-type.

Figure 10A is a schematic diagram illustrating the level of thermal stability (represented by T_m) improvement over “wild-type” obtained by GSSM™ evolution. The
10 single site amino acid mutant and the combinatorial variant with the highest thermal stability (Q11H and “9X” (SEQ ID NO:378), respectively) are shown in comparison to wild-type. Figure 10B illustrates a “fitness diagram” of enzyme improvement obtained by combining GSSM™ and GeneReassembly™ technologies. Fitness was determined using the formula $F = F_T + F_V$ where fitness (F) is calculated by equally weighting thermal tolerance fitness (F_T)
15 and relative activity fitness (F_V) as described above. The point mutation that confers the greatest fitness (S65V) is shown. Combining all 9 point mutations also improved fitness (SEQ ID NO:378, the “9X” variant). However, the largest improvement in fitness was obtained by combining GSSM™ and GeneReassembly™ methods to obtain the best variant, 6X-2 (SEQ ID NO:380).

20 The GeneReassembly™ method also allowed the identification of important residues that appear absolutely necessary for improved thermal stability. Two key residues, Q11H and G17I, were present in every GeneReassembly™ variant identified based on thermal tolerance (see Figure 6A). The structural determinants for thermal stability of proteins have been studied and several theories have been documented, e.g., by Kinjo (2001)
25 Eur. Biophys. J. 30:378-384; Britton (1999) J. Mol. Biol. 293:1121-1132; Ladenstein (1998) Adv. Biochem. Eng. Biotechnol. 61:37-85; Britton (1995) Eur. J. Biochem. 229:688-695; Tanner (1996) Biochemistry 35:2597-2609; Vetriani (1998) Proc. Natl. Acad. Sci. USA 95:2300-2305. Hydrogen bonding patterns, ionic interactions, hydrophobic packing and decreased length of surface loops are among the key factors even though the contribution of
30 each to protein stability is not fully understood. Given that most of the beneficial point substitutions identified from testing all possible single amino acid substitutions involved the replacement of relatively polar, charged or small (glycine) residues for much larger hydrophobic residues, it can be surmised that hydrophobic interactions play the most significant

role in enhancing the thermostability of this protein. Even with a good understanding of the optimal interactions to enhance thermal tolerance, the prediction of where to make mutations that introduce such interactions is not straightforward. A nonrational approach using the GSSM™ method, however, allows rapid sampling of all sidechains at all positions within a protein structure. Such an approach leads to the discovery of amino acid substitutions that introduce functional interactions that could not have been foreseen.

EXAMPLE 6: Pre-treating paper pulp with xylanases of the invention

In one aspect, xylanases of the invention can be used to pretreat paper pulp. This example describes an exemplary routine screening protocol to determine whether a xylanase is useful in pretreating paper pulp; e.g., in reducing the use of bleaching chemicals (e.g., chlorine dioxide, ClO₂) when used to pretreat Kraft paper pulp.

The screening protocol has two alternative test parameters: Impact of xylanase treatment after an oxygen delignification step (post-O₂ pulp); and, Impact of xylanase in a process that does not include oxygen delignification (pre-O₂ brownstock).

For pulp treatment conditions that simulate process conditions in industrial situations, e.g., factories: pH 8.0; 70 °C; 60 min duration.

The process is schematically depicted in the Flow Diagram of Figure 11.

Twenty xylanases were identified by biochemical tests that were active under these conditions. Of the 20 xylanases, 6 were able to significantly reduce ClO₂ demand when they were used to pretreat Kraft pulp before it was chemically bleached. The six are: SEQ ID NO:182 (encoded by SEQ ID NO:181); SEQ ID NO:160 (encoded by SEQ ID NO:159); SEQ ID NO:198 (encoded by SEQ ID NO:197); SEQ ID NO:168 (encoded by SEQ ID NO:167); SEQ ID NO:216 (encoded by SEQ ID NO:215); SEQ ID NO:260 (encoded by SEQ ID NO:259). Others showed some activity but were not as good. Xylanases SEQ ID NO:182 (encoded by SEQ ID NO:181) and SEQ ID NO:160 (encoded by SEQ ID NO:159) are modular and contain a carbohydrate binding module in addition to the xylanase catalytic domain. It was demonstrated that truncated derivatives of these 2 xylanases containing just the catalytic domain are more effective in this application. The best xylanase, SEQ ID NO:160 (encoded by SEQ ID NO:159) was studied more comprehensively. Results can be summarized as follows:

- pretreatment of post-O₂ spruce/pine/fir (SPF) pulp with 2 units/g of SEQ ID NO:160 (encoded by SEQ ID NO:159) reduces subsequent ClO₂ use by 22% to reach 65%GE brightness;

- pretreatment of pre-O₂ brownstock SPF with 0.5 units/g SEQ ID NO:160 (encoded by SEQ ID NO:159) reduces subsequent ClO₂ use by 13% to reach 65%GE brightness;
 - pretreatment of pre-O₂ Aspen pulp with 0.5 units/g SEQ ID NO:160 (encoded by SEQ ID NO:159) reduces ClO₂ use by at least 22%;
 - pretreatment of pre-O₂ Douglas Fir/Hemlock pulp with 0.5 units/g SEQ ID NO:160 (encoded by SEQ ID NO:159) reduces ClO₂ use by at least 22%;
 - under the treatment conditions employed, the reduction in yield from the xylanase treatment did not exceed 0.5% when compared with pulp that had been bleached at the same kappa factor but not treated with xylanase;
 - optimal conditions for treating post-O₂ SPF pulp with SEQ ID NOS:159, 160 were: pH 6-7, enzyme dose 0.3 units/g, treatment time 20-25 min. Under these conditions, reduction in ClO₂ use of 28% was possible to reach 69%GE brightness.
- In further experiments:
- XYLA (E.c) = truncated variant of SEQ ID NOS:159, 160 containing only xylanase catalytic domain expressed in *E.coli*
 - XYLA (P.f) = ditto but expressed in *P. fluorescens*
 - XYLB (E.c) = truncated variant etc, etc expressed in *E.coli*
 - XYLB (P.f) = ditto but expressed in *P. fluorescens*

Dose Response Data for Lead Xylanases on Pre-O₂ Brownstock

- Conditions for xylanase stage (X-stage) as follows:
- pH 8
 - Temperature 70°C
 - Time 60 min
 - Kappa factor 0.24
- For no-enzyme control, kappa factor was 0.30
- Results showed a dose dependent increase in brightness for xylanase-treated samples at a lower charge of chlorine dioxide (ClO₂) (Kf 0.24 vs Kf 0.30).

In each case, the truncated derivative looked to be more effective than the full-length xylanase. Optimal xylanase dose looked to be around 0.6 to 0.7 U/g pulp.

Pretreatment of Intercontinental Pre-O₂ Brownstock with the best 4 Xylanases

Determination of ClO₂ Dose Response in D₀

5 Experimental outline

- Pre-O₂ Brownstock
 - Initial kappa 31.5
- X stage conditions
 - Xylanase charge 0.7 U/gm
 - 10 ○ Temperature 70°C
 - pH 8
 - Treatment time 1 hr
 - Pulp consistency 10%
- Bleach sequence XDE_p
 - 15 ○ Kappa factor 0.22, 0.26 and 0.30 (%D on pulp: 2.63, 3.12 and 3.60)

Final brightness after 3-stage bleach sequence versus Kappa factor (ClO₂ charge):

- XYLB - At 61.5 final brightness, X-stage enables reduction in ClO₂ use of 3.89 kg/ton pulp.
- XYLB (E.c) - At 61.5 final brightness, X-stage enables reduction in ClO₂ charge of 4.07 kg/ton pulp.
- 20 • XYLA - At 61.5 brightness, X-stage enables a reduction in ClO₂ use of 4.07 kg/ton pulp.
- XYLA (E.c) - At 61.5 final brightness, X-stage enables reduction in ClO₂ use of 4.90 kg/ton pulp.

25 Determination of ClO₂ Dose Response in D₀:

Enzyme	ClO ₂ Savings in D ₀ (kg/ton OD)	Kf reduction in D ₀
XYLB	3.89	11.7%
XYLB (E.c)	5.08	15.8%
XYLA	4.07	12.2%
XYLA (E.c)	4.90	14.7%

Xylanase 0.7 U/g, pH 8.0, 70 °C, 1 hr

Pulp: Pre-O₂ Brownstock, initial kappa 31.5

Percentage saving of ClO₂ is of little significance to the industry. Their
 5 primary concern is lbs of ClO₂ required per ton OD pulp. This makes sense when one
 considers that a lower percentage saving seen with a high initial kappa brownstock can be
 more valuable in terms of lbs of ClO₂ saved than a higher percentage reduction for a low
 initial kappa pulp which will require a lower total charge of ClO₂ to reach target brightness.

Relationship between Brightness, Yield and Kappa Factor for Bleached Control Pulp:

10 The results showed that bleaching with increasing doses of ClO₂ to achieve
 higher target brightness results in increased loss of pulp yield. This is an issue because pulp at
 this stage of the process has a value of almost \$400 per ton and loss of cellulose costs money.

A benefit of xylanase (e.g., a xylanase of the invention) is that use of a lower
 ClO₂ dose can reduce yield losses as long as the action of the xylanase itself doesn't cancel
 15 out the gain.

Dose Response Data for Pretreatment of Pre-O₂ Brownstock with Xylanase XYLB (P.f):

Experimental outline

- Northwood Pre-O₂ Brownstock
- Initial kappa 28.0
- 20 –Initial consistency 32.46%
- Initial brightness 28.37
- X stage conditions
- Xylanase charge 0 to 2.70 U/gm
- Temperature 58°C to 61°C

–pH 8.2 to 8.5

–Treatment time 1hr

•Bleach sequence XDE_p

–Kappa factor 0.24

- 5 •ClO₂ saving calculated for Kappa factors between 0.24 and 0.30

The purpose of this experiment was to evaluate the best of the 4 xylanases on unwashed SPF brownstock. Results showed dose-dependent increases in final brightness for pulp treated with XYLB (E.c), with brightness achieved in presence of xylanase at lower Kf of 0.24, approaching brightness achieved at higher Kf of 0.30 asymptotically.

- 10 Relationship between Dose of Xylanase XYLB (E.c) and Chlorine Dioxide Saving (Pre-O₂ Brownstock):

ClO ₂ Saving in % OD Pulp	ClO ₂ Saving in kg/ton Pulp	Xylanase Dose in U/gm
0.299%	2.99	0.31
0.363%	3.63	0.51
0.406%	4.06	0.71
0.439%	4.39	0.91
0.483%	4.83	1.26
0.523%	5.32	1.80
0.587%	5.87	2.70

Optimum Xylanase Dose is between 0.5 and 0.9 U/gm

- 15 The optimum dose lies in the range 0.5 to 0.9 U/g. Above this dose there is a diminishing return per unit increment of xylanase. Reductions in chlorine dioxide dose per ton of pulp treated of this magnitude are commercially significant.

Three-stage biobleaching procedure

- 20 A three-stage biobleaching procedure was developed that would closely simulate the actual bleaching operations in a pulp mill bleach plant (Fig. 1). This bleach sequence is designated by (X)DoEp, in which X represents the xylanase treatment stage, D for chlorine dioxide bleaching stage, and E_p for alkaline peroxide extraction stage. The primary feedstock used in our application tests was Southern Softwood Kraft Brownstock (without oxygen delignification). The most effective xylanase candidates that showed high

bleach chemical reduction potential in the biobleaching assays were also tested on two species of hardwood Kraft pulp (maple and aspen). Upon completion of each biobleaching round, the ensuing pulp was used to produce TAPPI (Technical Association of Pulp and Paper Industries)-standard handsheets. The GE% brightness of each handsheet was
5 measured, and the brightness values were used as the indication of how well each enzyme had performed on the pulp during the enzymatic pretreatment stage (X).

Results:

Out of approximately 110 xylanases that were screened using the (X)DoEp biobleaching sequence, 4 enzymes, i.e., XYLA (P.f); XYLB (P.f); SEQ ID NO216 (encoded
10 by SEQ ID NO:215); SEQ ID NO:176 (encoded by SEQ ID NO: 175); showed the greatest potential for reducing the use of bleaching chemicals. While XYLA (P.f) and XYLB (P.f) exhibited equally high performance (best among the four good performers), XYLA (P.f) showed a better pH tolerance than XYLB (P.f). The results can be summarized as follows:

- It is possible to achieve a handsheet brightness of 60 (GE%) using a three-stage bleach
15 sequence [(X)DoEp] that involves pretreatment of Southern Softwood Kraft Brownstock with the following four enzymes at the loading levels listed below (pH=8, 65 °C & 1 h):
 - XYLA (P.f) at 0.55 U/g pulp
 - XYLB (P.f) at 0.75 U/g pulp
 - SEQ ID NOS:215, 216 at 1.80 U/g pulp
 - 20 ○ SEQ ID NOS:175, 176 at 1.98 U/g pulp
- Pretreatment of Southern Softwood Kraft Brownstock with 2 U/g pulp of XYLA (P.f) reduces ClO₂ use by 18.7% to reach a final GE% brightness of 61.
- XYLA (P.f) exhibits good tolerance at higher pH and provides more than 14% chemical savings when the enzymatic pretreatment stage is run at pH=10.
- 25 • Pretreatment of Southern Softwood Kraft Brownstock with 2 U/g pulp of XYLB (P.f) reduces ClO₂ use by 16.3% to reach a final GE% brightness of 60.5.
- Pretreatment of aspen Kraft pulp with 2 U/g pulp of XYLA (P.f) and XYLB (P.f) reduces ClO₂ use by about 35% to reach a final GE% brightness of 77.
- Pretreatment of maple Kraft pulp with 2 U/g pulp of XYLA (P.f) and XYLB (P.f)
30 reduces ClO₂ use by about 38% to reach a final GE% brightness of 79.

- The two best performing xylanases, namely XYLA (P.f) and XYLB (P.f), are truncated enzymes, containing just the catalytic domain, and were produced in *Pseudomonas fluorescens*.

5 While the invention has been described in detail with reference to certain preferred aspects thereof, it will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

WHAT IS CLAIMED IS:

1. An isolated or recombinant nucleic acid comprising a nucleic acid
5 sequence having at least 50% sequence identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID
NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ
ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27,
SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID
NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49,
10 SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID
NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71,
SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID
NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93,
SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID
15 NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID
NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID
NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID
NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID
NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID
20 NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID
NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID
NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID
NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID
NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID
25 NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID
NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID
NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID
NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID
NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID
30 NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID
NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID
NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID
NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID
NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID

NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, over a region of at least about 100 residues, wherein the nucleic acid encodes at least one polypeptide having a xylanase activity, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection.

2. The isolated or recombinant nucleic acid of claim 1, wherein the sequence identity is at least about 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63% or 64%.

3. The isolated or recombinant nucleic acid of claim 1, wherein the sequence identity is at least about 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID

NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID
 NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID
 NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID
 NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID
 5 NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID
 NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID
 NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID
 NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID
 NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID
 10 NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID
 NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID
 NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID
 NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID
 NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID
 15 NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID
 NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID
 NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID
 NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID
 NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID
 20 NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID
 NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID
 NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID
 NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379.

25 4. The isolated or recombinant nucleic acid of claim 1, wherein the
 sequence identity is over a region of at least about 50, 75, 100, 150, 200, 250, 300, 350, 400,
 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150 or more
 residues, or the full length of a gene or a transcript.

30 5. The isolated or recombinant nucleic acid of claim 1, wherein the
 nucleic acid sequence comprises a sequence as set forth in SEQ ID NO:1, SEQ ID NO:3,
 SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID
 NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25,
 SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID

NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47,
SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID
NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69,
SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID
5 NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91,
SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID
NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID
NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID
NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID
10 NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID
NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID
NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID
NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID
NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID
15 NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID
NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID
NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID
NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID
NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID
20 NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID
NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID
NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID
NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID
NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID
25 NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID
NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID
NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID
NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID
NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID
30 NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID
NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID
NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID
NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID
NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379.

6. The isolated or recombinant nucleic acid of claim 1, wherein the nucleic acid sequence encodes a polypeptide having a sequence as set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198, SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228, SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:244, SEQ ID NO:246, SEQ ID NO:248, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256, SEQ ID NO:258, SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:268, SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276, SEQ ID NO:278, SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286, SEQ ID NO:288, SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296, SEQ ID NO:298, SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306, SEQ ID NO:308, SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:322,

SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332,
SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342,
SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352,
SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362,
5 SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372,
SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378 or SEQ ID NO:380.

7. The isolated or recombinant nucleic acid of claim 1, wherein the
sequence comparison algorithm is a BLAST version 2.2.2 algorithm where a filtering setting
10 is set to blastall -p blastp -d "nr pataa" -F F, and all other options are set to default.

8. The isolated or recombinant nucleic acid of claim 1, wherein the
xylanase activity comprises catalyzing hydrolysis of internal β -1,4-xylosidic linkages.

15 9. The isolated or recombinant nucleic acid of claim 8, wherein the
xylanase activity comprises an endo-1,4-beta-xylanase activity.

10. The isolated or recombinant nucleic acid of claim 1, wherein the
xylanase activity comprises hydrolyzing a xylan to produce a smaller molecular weight
20 xylose and xylo-oligomer.

11. The isolated or recombinant nucleic acid of claim 10, wherein the
xylan comprises an arabinoxylan.

25 12. The isolated or recombinant nucleic acid of claim 11, wherein the
arabinoxylan comprises a water soluble arabinoxylan.

13. The isolated or recombinant nucleic acid of claim 12, wherein the
water soluble arabinoxylan comprises a dough or a bread product.

30 14. The isolated or recombinant nucleic acid of claim 1, wherein the
xylanase activity comprises hydrolyzing polysaccharides comprising 1,4- β -glycoside-linked
D-xylopyranoses.

15. The isolated or recombinant nucleic acid of claim 1, wherein the xylanase activity comprises hydrolyzing hemicelluloses.
16. The isolated or recombinant nucleic acid of claim 15, wherein the xylanase activity comprises hydrolyzing hemicelluloses in a wood or paper pulp or a paper product.
17. The isolated or recombinant nucleic acid of claim 8, wherein the xylanase activity comprises catalyzing hydrolysis of xylans in a feed or a food product.
18. The isolated or recombinant nucleic acid of claim 17, wherein the feed or food product comprises a cereal-based animal feed, a wort or a beer, a milk or a milk product, a fruit or a vegetable.
19. The isolated or recombinant nucleic acid of claim 1, wherein the xylanase activity comprises catalyzing hydrolysis of xylans in a microbial cell or a plant cell.
20. The isolated or recombinant nucleic acid of claim 1, wherein the xylanase activity is thermostable.
21. The isolated or recombinant nucleic acid of claim 20, wherein the polypeptide retains a xylanase activity under conditions comprising a temperature range of between about 37°C to about 95°C, or between about 55°C to about 85°C, or between about 70°C to about 75°C, or between about 70°C to about 95°C, or between about 90°C to about 95°C.
22. The isolated or recombinant nucleic acid of claim 1, wherein the xylanase activity is thermotolerant.
23. The isolated or recombinant nucleic acid of claim 22, wherein the polypeptide retains a xylanase activity after exposure to a temperature in the range from greater than 37°C to about 95°C, from greater than 55°C to about 85°C, or between about 70°C to about 75°C, or from greater than 90°C to about 95°C.

24. An isolated or recombinant nucleic acid, wherein the nucleic acid comprises a sequence that hybridizes under stringent conditions to a nucleic acid comprising SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID
5 NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77,
10 SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129,
15 SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179,
20 SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229,
25 SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279,
30 SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329,

SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339,
 SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349,
 SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359,
 SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369,
 5 SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379,
 wherein the nucleic acid encodes a polypeptide having a xylanase activity.

25. The isolated or recombinant nucleic acid of claim 24, wherein the
 nucleic acid is at least about 50, 75, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000 or
 10 more residues in length or the full length of the gene or transcript.

26. The isolated or recombinant nucleic acid of claim 24, wherein the
 stringent conditions include a wash step comprising a wash in 0.2X SSC at a temperature of
 about 65°C for about 15 minutes.

15 27. A nucleic acid probe for identifying a nucleic acid encoding a
 polypeptide with a xylanase activity, wherein the probe comprises at least 10 consecutive
 bases of a sequence comprising SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7,
 SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID
 20 NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29,
 SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID
 NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51,
 SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID
 NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73,
 25 SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID
 NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95,
 SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ
 ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID
 NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID
 30 NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID
 NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID
 NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID
 NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID
 NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID

NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, wherein the probe identifies the nucleic acid by binding or hybridization.

28. The nucleic acid probe of claim 27, wherein the probe comprises an oligonucleotide comprising at least about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, about 60 to 100, or about 50 to 150 consecutive bases.

29. A nucleic acid probe for identifying a nucleic acid encoding a polypeptide having a xylanase activity, wherein the probe comprises a nucleic acid comprising at least about 10 consecutive residues of a nucleic acid sequence having at least 50% sequence identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41,

SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID
NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63,
SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID
NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85,
5 SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID
NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID
NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID
NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID
NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID
10 NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID
NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID
NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID
NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID
NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID
15 NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID
NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID
NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID
NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID
NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID
20 NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID
NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID
NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID
NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID
NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID
25 NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID
NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID
NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID
NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID
NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID
30 NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID
NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID
NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID
NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID

NO:377 or SEQ ID NO:379, wherein the sequence identities are determined by analysis with a sequence comparison algorithm or by visual inspection.

30. The nucleic acid probe of claim 29, wherein the probe comprises an
5 oligonucleotide comprising at least about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, about 60 to 100, or about 50 to 150 consecutive bases.

31. An amplification primer pair for amplifying a nucleic acid encoding a
polypeptide having a xylanase activity, wherein the primer pair is capable of amplifying a
10 nucleic acid comprising a sequence as set forth in claim 1 or claim 24, or a subsequence thereof.

32. The amplification primer pair of claim 31, wherein a member of the
amplification primer sequence pair comprises an oligonucleotide comprising at least about 10
15 to 50 consecutive bases of the sequence, or, about 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more consecutive bases of the sequence.

33. An amplification primer pair, wherein the primer pair comprises a first
member having a sequence as set forth by about the first (the 5') 12, 13, 14, 15, 16, 17, 18,
20 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or more residues of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47,
25 SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID

NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID
NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID
NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID
NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID
5 NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID
NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID
NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID
NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID
NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID
10 NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID
NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID
NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID
NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID
NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID
15 NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID
NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID
NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID
NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID
NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID
20 NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID
NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID
NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID
NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, and a second member
having a sequence as set forth by about the first (the 5') 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,
25 22, 23, 24, 25, 26, 27, 28, 29, 30 or more residues of the complementary strand of the first
member.

34. A xylanase-encoding nucleic acid generated by amplification of a
polynucleotide using an amplification primer pair as set forth in claim 33.

30

35. The xylanase-encoding nucleic acid of claim 34, wherein the
amplification is by polymerase chain reaction (PCR).

36. The xylanase-encoding nucleic acid of claim 34, wherein the nucleic acid generated by amplification of a gene library.

37. The xylanase-encoding nucleic acid of claim 34, wherein the gene library is an environmental library.

38. An isolated or recombinant xylanase encoded by a xylanase-encoding nucleic acid as set forth in claim 34.

39. A method of amplifying a nucleic acid encoding a polypeptide having a xylanase activity comprising amplification of a template nucleic acid with an amplification primer sequence pair capable of amplifying a nucleic acid sequence as set forth in claim 1 or claim 24, or a subsequence thereof.

40. An expression cassette comprising a nucleic acid comprising a sequence as set forth in claim 1 or claim 24.

41. A vector comprising a nucleic acid comprising a sequence as set forth in claim 1 or claim 24.

42. A cloning vehicle comprising a nucleic acid comprising a sequence as set forth in claim 1 or claim 24, wherein the cloning vehicle comprises a viral vector, a plasmid, a phage, a phagemid, a cosmid, a fosmid, a bacteriophage or an artificial chromosome.

43. The cloning vehicle of claim 42, wherein the viral vector comprises an adenovirus vector, a retroviral vector or an adeno-associated viral vector.

44. The cloning vehicle of claim 42, comprising a bacterial artificial chromosome (BAC), a plasmid, a bacteriophage P1-derived vector (PAC), a yeast artificial chromosome (YAC), or a mammalian artificial chromosome (MAC).

45. A transformed cell comprising a nucleic acid comprising a sequence as set forth in claim 1 or claim 24.

46. A transformed cell comprising an expression cassette as set forth in claim 40.

5 47. The transformed cell of claim 40, wherein the cell is a bacterial cell, a mammalian cell, a fungal cell, a yeast cell, an insect cell or a plant cell.

48. A transgenic non-human animal comprising a sequence as set forth in claim 1 or claim 24.

10

49. The transgenic non-human animal of claim 48, wherein the animal is a mouse.

15

50. A transgenic plant comprising a sequence as set forth in claim 1 or claim 24.

20

51. The transgenic plant of claim 50, wherein the plant is a corn plant, a sorghum plant, a potato plant, a tomato plant, a wheat plant, an oilseed plant, a rapeseed plant, a soybean plant, a rice plant, a barley plant, a grass, or a tobacco plant.

52. A transgenic seed comprising a sequence as set forth in claim 1 or claim 24.

25

53. The transgenic seed of claim 52, wherein the seed is a corn seed, a wheat kernel, an oilseed, a rapeseed, a soybean seed, a palm kernel, a sunflower seed, a sesame seed, a rice, a barley, a peanut or a tobacco plant seed.

30

54. An antisense oligonucleotide comprising a nucleic acid sequence complementary to or capable of hybridizing under stringent conditions to a sequence as set forth in claim 1 or claim 24, or a subsequence thereof.

55. The antisense oligonucleotide of claim 49, wherein the antisense oligonucleotide is between about 10 to 50, about 20 to 60, about 30 to 70, about 40 to 80, or about 60 to 100 bases in length.

56. A method of inhibiting the translation of a xylanase message in a cell comprising administering to the cell or expressing in the cell an antisense oligonucleotide comprising a nucleic acid sequence complementary to or capable of hybridizing under
5 stringent conditions to a sequence as set forth in claim 1 or claim 24.

57. A double-stranded inhibitory RNA (RNAi) molecule comprising a subsequence of a sequence as set forth in claim 1 or claim 24.

10 58. The double-stranded inhibitory RNA (RNAi) molecule of claim 52, wherein the RNAi is about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 or more duplex nucleotides in length.

59. A method of inhibiting the expression of a xylanase in a cell
15 comprising administering to the cell or expressing in the cell a double-stranded inhibitory RNA (iRNA), wherein the RNA comprises a subsequence of a sequence as set forth in claim 1 or claim 24.

60. An isolated or recombinant polypeptide (i) having at least 50%
20 sequence identity to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID
25 NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID
30 NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NO:138, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158,

SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168,
SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178,
SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188,
SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198,
5 SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208,
SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218,
SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228,
SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238,
SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:244, SEQ ID NO:246, SEQ ID NO:248,
10 SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256, SEQ ID NO:258,
SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:268,
SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276, SEQ ID NO:278,
SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286, SEQ ID NO:288,
SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296, SEQ ID NO:298,
15 SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306, SEQ ID NO:308,
SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316, SEQ ID NO:318,
SEQ ID NO:320, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328,
SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:338,
SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348,
20 SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358,
SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368,
SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378 or
SEQ ID NO:380, over a region of at least about 100 residues, wherein the sequence identities
are determined by analysis with a sequence comparison algorithm or by a visual inspection,
25 or, (ii) encoded by a nucleic acid having at least 50% sequence identity to a sequence as set
forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID
NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21,
SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID
NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43,
30 SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID
NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65,
SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID
NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87,
SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID

NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, over a region of at least about 100 residues, and the sequence identities are determined by analysis with a sequence comparison algorithm or by a visual inspection, or encoded by a nucleic acid capable of hybridizing under stringent conditions to a sequence as set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID

NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43,
SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID
NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65,
SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID
5 NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87,
SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID
NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID
NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID
NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID
10 NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID
NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID
NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID
NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID
NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID
15 NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID
NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID
NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID
NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID
NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID
20 NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID
NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID
NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID
NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID
NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID
25 NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID
NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID
NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID
NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID
NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID
30 NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID
NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID
NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID
NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID

NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379.

61. The isolated or recombinant polypeptide of claim 60, wherein the
5 sequence identity is over a region of at least about 51%, 52%, 53%, 54%, 55%,
56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%,
72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%,
88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or is 100%
sequence identity.

10

62. The isolated or recombinant polypeptide of claim 60, wherein the
sequence identity is over a region of at least about 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100,
150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000,
1050 or more residues, or the full length of an enzyme.

15

63. The isolated or recombinant polypeptide of claim 60, wherein the
polypeptide has a sequence as set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ
ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18,
SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID
20 NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40,
SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID
NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62,
SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID
NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84,
25 SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID
NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID
NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID
NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID
NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132; SEQ ID NO:134; SEQ ID
30 NO:136; SEQ ID NO:138; SEQ ID NO:140; SEQ ID NO:142; SEQ ID NO:144; NO:146,
SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156,
SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166,
SEQ ID NO:168, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176,
SEQ ID NO:178, SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186,

SEQ ID NO:188, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196,
SEQ ID NO:198, SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206,
SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216,
SEQ ID NO:218, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226,
5 SEQ ID NO:228, SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236,
SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:244, SEQ ID NO:246,
SEQ ID NO:248, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256,
SEQ ID NO:258, SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266,
SEQ ID NO:268, SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276,
10 SEQ ID NO:278, SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286,
SEQ ID NO:288, SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296,
SEQ ID NO:298, SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306,
SEQ ID NO:308, SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316,
SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326,
15 SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334, SEQ ID NO:336,
SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346,
SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356,
SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366,
SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376,
20 SEQ ID NO:378 or SEQ ID NO:380.

64. The isolated or recombinant polypeptide of claim 60, wherein the polypeptide has a xylanase activity.

25 65. The isolated or recombinant polypeptide of claim 64, wherein the xylanase activity comprises catalyzing hydrolysis of internal β -1,4-xylosidic linkages.

66. The isolated or recombinant polypeptide of claim 65, wherein the xylanase activity comprises an endo-1,4-beta-xylanase activity.

30 67. The isolated or recombinant polypeptide of claim 64, wherein the xylanase activity comprises hydrolyzing a xylan to produce a smaller molecular weight xylose and xylo-oligomer.

68. The isolated or recombinant polypeptide of claim 67, wherein the xylan comprises an arabinoxylan.

69. The isolated or recombinant polypeptide of claim 68, wherein the
5 arabinoxylan comprises a water soluble arabinoxylan.

70. The isolated or recombinant polypeptide of claim 69, wherein the water soluble arabinoxylan comprises a dough or a bread product.

10 71. The isolated or recombinant polypeptide of claim 64, wherein the xylanase activity comprises hydrolyzing polysaccharides comprising 1,4- β -glycoside-linked D-xylopyranoses.

72. The isolated or recombinant polypeptide of claim 64, wherein the
15 xylanase activity comprises hydrolyzing hemicelluloses.

73. The isolated or recombinant polypeptide of claim 72, wherein the xylanase activity comprises hydrolyzing hemicelluloses in a wood or paper pulp or a paper product.

20 74. The isolated or recombinant polypeptide of claim 73, wherein the xylanase activity comprises catalyzing hydrolysis of xylans in a feed or a food product.

75. The isolated or recombinant polypeptide of claim 74, wherein the feed
25 or food product comprises a cereal-based animal feed, a wort or a beer, a milk or a milk product, a fruit or a vegetable.

76. The isolated or recombinant polypeptide of claim 64, wherein the xylanase activity comprises catalyzing hydrolysis of xylans in a microbial cell or a plant cell.

30 77. The isolated or recombinant polypeptide of claim 64, wherein the xylanase activity is thermostable.

78. The isolated or recombinant polypeptide of claim 77, wherein the polypeptide retains a xylanase activity under conditions comprising a temperature range of between about 1°C to about 5°C, between about 5°C to about 15°C, between about 15°C to about 25°C, between about 25°C to about 37°C, between about 37°C to about 95°C, between
5 about 55°C to about 85°C, between about 70°C to about 95°C, between about 70°C to about 75°C, or between about 90°C to about 95°C.

79. The isolated or recombinant polypeptide of claim 64, wherein the xylanase activity is thermotolerant.

10

80. The isolated or recombinant polypeptide of claim 79, wherein the polypeptide retains a xylanase activity after exposure to a temperature in the range from between about 1°C to about 5°C, between about 5°C to about 15°C, between about 15°C to about 25°C, between about 25°C to about 37°C, between about 37°C to about 95°C, between
15 about 55°C to about 85°C, between about 70°C to about 75°C, or between about 90°C to about 95°C, or more.

81. An isolated or recombinant polypeptide comprising a polypeptide as set forth in claim 60 and lacking a signal sequence or a prepro sequence.

20

82. An isolated or recombinant polypeptide comprising a polypeptide as set forth in claim 60 and having a heterologous signal sequence or a heterologous prepro sequence.

83. The isolated or recombinant polypeptide of claim 64, wherein the xylanase activity comprises a specific activity at about 37°C in the range from about 100 to about 1000 units per milligram of protein, from about 500 to about 750 units per milligram of protein, from about 500 to about 1200 units per milligram of protein, or from about 750 to about 1000 units per milligram of protein.

30

84. The isolated or recombinant polypeptide of claim 79, wherein the thermotolerance comprises retention of at least half of the specific activity of the xylanase at 37°C after being heated to an elevated temperature.

85. The isolated or recombinant polypeptide of claim 79, wherein the thermotolerance comprises retention of specific activity at 37°C in the range from about 500 to about 1200 units per milligram of protein after being heated to an elevated temperature.

5 86. The isolated or recombinant polypeptide of claim 60, wherein the polypeptide comprises at least one glycosylation site.

87. The isolated or recombinant polypeptide of claim 86, wherein the glycosylation is an N-linked glycosylation.

10

88. The isolated or recombinant polypeptide of claim 87, wherein the polypeptide is glycosylated after being expressed in a *P. pastoris* or a *S. pombe*.

89. The isolated or recombinant polypeptide of claim 64, wherein the polypeptide retains a xylanase activity under conditions comprising about pH 6.5, pH 6.0, pH 5.5, 5.0, pH 4.5 or 4.0.

15

90. The isolated or recombinant polypeptide of claim 64, wherein the polypeptide retains a xylanase activity under conditions comprising about pH 7.5, pH 8.0, pH 8.5, pH 9, pH 9.5, pH 10 or pH 10.5.

20

91. A protein preparation comprising a polypeptide as set forth in claim 60, wherein the protein preparation comprises a liquid, a solid or a gel.

92. A heterodimer comprising a polypeptide as set forth in claim 60 and a second domain.

25

93. The heterodimer of claim 92, wherein the second domain is a polypeptide and the heterodimer is a fusion protein.

30

94. The heterodimer of claim 92, wherein the second domain is an epitope or a tag.

95. A homodimer comprising a polypeptide as set forth in claim 60.

96. An immobilized polypeptide, wherein the polypeptide comprises a sequence as set forth in claim 60, or a subsequence thereof.

5 97. The immobilized polypeptide of claim 96, wherein the polypeptide is immobilized on a cell, a metal, a resin, a polymer, a ceramic, a glass, a microelectrode, a graphitic particle, a bead, a gel, a plate, an array or a capillary tube.

98. An array comprising an immobilized polypeptide as set forth in claim
10 60.

99. An array comprising an immobilized nucleic acid as set forth in claim 1 or claim 24.

15 100. An isolated or recombinant antibody that specifically binds to a polypeptide as set forth in claim 60.

101. The isolated or recombinant antibody of claim 100, wherein the antibody is a monoclonal or a polyclonal antibody.

20 102. A hybridoma comprising an antibody that specifically binds to a polypeptide as set forth in claim 60.

103. A method of isolating or identifying a polypeptide with a xylanase
25 activity comprising the steps of:

(a) providing an antibody as set forth in claim 100;
(b) providing a sample comprising polypeptides; and
(c) contacting the sample of step (b) with the antibody of step (a) under conditions wherein the antibody can specifically bind to the polypeptide, thereby isolating or
30 identifying a polypeptide having a xylanase activity.

104. A method of making an anti-xylanase antibody comprising administering to a non-human animal a nucleic acid as set forth in claim 1 or claim 24 or a

subsequence thereof in an amount sufficient to generate a humoral immune response, thereby making an anti-xylanase antibody.

105. A method of making an anti-xylanase antibody comprising
5 administering to a non-human animal a polypeptide as set forth in claim 60 or a subsequence thereof in an amount sufficient to generate a humoral immune response, thereby making an anti-xylanase antibody.

106. A method of producing a recombinant polypeptide comprising the
10 steps of: (a) providing a nucleic acid operably linked to a promoter, wherein the nucleic acid comprises a sequence as set forth in claim 1 or claim 24; and (b) expressing the nucleic acid of step (a) under conditions that allow expression of the polypeptide, thereby producing a recombinant polypeptide.

15 107. The method of claim 106, further comprising transforming a host cell with the nucleic acid of step (a) followed by expressing the nucleic acid of step (a), thereby producing a recombinant polypeptide in a transformed cell.

108. A method for identifying a polypeptide having a xylanase activity
20 comprising the following steps:
(a) providing a polypeptide as set forth in claim 64;
(b) providing a xylanase substrate; and
(c) contacting the polypeptide with the substrate of step (b) and detecting a decrease in the amount of substrate or an increase in the amount of a reaction product,
25 wherein a decrease in the amount of the substrate or an increase in the amount of the reaction product detects a polypeptide having a xylanase activity.

109. A method for identifying a xylanase substrate comprising the
following steps:
30 (a) providing a polypeptide as set forth in claim 64;
(b) providing a test substrate; and
(c) contacting the polypeptide of step (a) with the test substrate of step (b) and detecting a decrease in the amount of substrate or an increase in the amount of reaction

product, wherein a decrease in the amount of the substrate or an increase in the amount of a reaction product identifies the test substrate as a xylanase substrate.

110. A method of determining whether a test compound specifically binds
5 to a polypeptide comprising the following steps:
- (a) expressing a nucleic acid or a vector comprising the nucleic acid under conditions permissive for translation of the nucleic acid to a polypeptide, wherein the nucleic acid has a sequence as set forth in claim 1 or claim 24;
 - (b) providing a test compound;
 - 10 (c) contacting the polypeptide with the test compound; and
 - (d) determining whether the test compound of step (b) specifically binds to the polypeptide.

111. A method of determining whether a test compound specifically binds
15 to a polypeptide comprising the following steps:
- (a) providing a polypeptide as set forth in claim 60;
 - (b) providing a test compound;
 - (c) contacting the polypeptide with the test compound; and
 - (d) determining whether the test compound of step (b) specifically binds to the
20 polypeptide.

112. A method for identifying a modulator of a xylanase activity comprising the following steps:
- (a) providing a polypeptide as set forth in claim 64;
 - 25 (b) providing a test compound;
 - (c) contacting the polypeptide of step (a) with the test compound of step (b) and measuring an activity of the xylanase, wherein a change in the xylanase activity measured in the presence of the test compound compared to the activity in the absence of the test compound provides a determination that the test compound modulates the xylanase
30 activity.

113. The method of claim 112, wherein the xylanase activity is measured by providing a xylanase substrate and detecting a decrease in the amount of the substrate or an

increase in the amount of a reaction product, or, an increase in the amount of the substrate or a decrease in the amount of a reaction product.

114. The method of claim 113, wherein a decrease in the amount of the
5 substrate or an increase in the amount of the reaction product with the test compound as compared to the amount of substrate or reaction product without the test compound identifies the test compound as an activator of a xylanase activity.

115. The method of claim 113, wherein an increase in the amount of the
10 substrate or a decrease in the amount of the reaction product with the test compound as compared to the amount of substrate or reaction product without the test compound identifies the test compound as an inhibitor of a xylanase activity.

116. A computer system comprising a processor and a data storage device
15 wherein said data storage device has stored thereon a polypeptide sequence or a nucleic acid sequence, wherein the polypeptide sequence comprises sequence as set forth in claim 60, a polypeptide encoded by a nucleic acid as set forth in claim 1 or claim 24.

117. The computer system of claim 115, further comprising a sequence
20 comparison algorithm and a data storage device having at least one reference sequence stored thereon.

118. The computer system of claim 117, wherein the sequence comparison
algorithm comprises a computer program that indicates polymorphisms.

25 119. The computer system of claim 117, further comprising an identifier that identifies one or more features in said sequence.

120. A computer readable medium having stored thereon a polypeptide
30 sequence or a nucleic acid sequence, wherein the polypeptide sequence comprises a polypeptide as set forth in claim 60; a polypeptide encoded by a nucleic acid as set forth in claim 1 or claim 24.

121. A method for identifying a feature in a sequence comprising the steps of: (a) reading the sequence using a computer program which identifies one or more features in a sequence, wherein the sequence comprises a polypeptide sequence or a nucleic acid sequence, wherein the polypeptide sequence comprises a polypeptide as set forth in claim 60; 5 a polypeptide encoded by a nucleic acid as set forth in claim 1 or claim 24; and (b) identifying one or more features in the sequence with the computer program.

122. A method for comparing a first sequence to a second sequence comprising the steps of: (a) reading the first sequence and the second sequence through use 10 of a computer program which compares sequences, wherein the first sequence comprises a polypeptide sequence or a nucleic acid sequence, wherein the polypeptide sequence comprises a polypeptide as set forth in claim 60 or a polypeptide encoded by a nucleic acid as set forth in claim 1 or claim 24; and (b) determining differences between the first sequence and the second sequence with the computer program.

123. The method of claim 122, wherein the step of determining differences between the first sequence and the second sequence further comprises the step of identifying polymorphisms.

124. The method of claim 123, further comprising an identifier that 20 identifies one or more features in a sequence.

125. The method of claim 124, comprising reading the first sequence using a computer program and identifying one or more features in the sequence.

126. A method for isolating or recovering a nucleic acid encoding a polypeptide with a xylanase activity from an environmental sample comprising the steps of: 25 (a) providing an amplification primer sequence pair as set forth in claim 31 or claim 33; (b) isolating a nucleic acid from the environmental sample or treating the environmental sample such that nucleic acid in the sample is accessible for hybridization to the amplification primer pair; and, 30

(c) combining the nucleic acid of step (b) with the amplification primer pair of step (a) and amplifying nucleic acid from the environmental sample, thereby isolating or

recovering a nucleic acid encoding a polypeptide with a xylanase activity from an environmental sample.

127. The method of claim 126, wherein each member of the amplification
5 primer sequence pair comprises an oligonucleotide comprising at least about 10 to 50
consecutive bases of a sequence as set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5,
SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID
NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27,
SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID
10 NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49,
SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID
NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71,
SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID
NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93,
15 SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID
NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID
NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID
NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID
NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID
20 NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID
NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID
NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID
NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID
NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID
25 NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID
NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID
NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID
NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID
NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID
30 NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID
NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID
NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID
NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID
NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID

NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, or a subsequence thereof.

10

128. A method for isolating or recovering a nucleic acid encoding a polypeptide with a xylanase activity from an environmental sample comprising the steps of:

(a) providing a polynucleotide probe comprising a sequence as set forth in claim 1 or claim 24, or a subsequence thereof;

15

(b) isolating a nucleic acid from the environmental sample or treating the environmental sample such that nucleic acid in the sample is accessible for hybridization to a polynucleotide probe of step (a);

(c) combining the isolated nucleic acid or the treated environmental sample of step (b) with the polynucleotide probe of step (a); and

20

(d) isolating a nucleic acid that specifically hybridizes with the polynucleotide probe of step (a), thereby isolating or recovering a nucleic acid encoding a polypeptide with a xylanase activity from an environmental sample.

25

129. The method of claim 127 or claim 128, wherein the environmental sample comprises a water sample, a liquid sample, a soil sample, an air sample or a biological sample.

30

130. The method of claim 129, wherein the biological sample is derived from a bacterial cell, a protozoan cell, an insect cell, a yeast cell, a plant cell, a fungal cell or a mammalian cell.

131. A method of generating a variant of a nucleic acid encoding a polypeptide with a xylanase activity comprising the steps of:

(a) providing a template nucleic acid comprising a sequence as set forth in claim 1 or claim 24; and

(b) modifying, deleting or adding one or more nucleotides in the template sequence, or a combination thereof, to generate a variant of the template nucleic acid.

5

132. The method of claim 131, further comprising expressing the variant nucleic acid to generate a variant xylanase polypeptide.

133. The method of claim 131, wherein the modifications, additions or deletions are introduced by a method comprising error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, gene site saturated mutagenesis (GSSM™), synthetic ligation reassembly (SLR) and a combination thereof.

15

134. The method of claim 131, wherein the modifications, additions or deletions are introduced by a method comprising recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation and a combination thereof.

135. The method of claim 131, wherein the method is iteratively repeated until a xylanase having an altered or different activity or an altered or different stability from that of a polypeptide encoded by the template nucleic acid is produced.

136. The method of claim 135, wherein the variant xylanase polypeptide is thermotolerant, and retains some activity after being exposed to an elevated temperature.

137. The method of claim 135, wherein the variant xylanase polypeptide has increased glycosylation as compared to the xylanase encoded by a template nucleic acid.

138. The method of claim 135, wherein the variant xylanase polypeptide has a xylanase activity under a high temperature, wherein the xylanase encoded by the template nucleic acid is not active under the high temperature.

5 139. The method of claim 131, wherein the method is iteratively repeated until a xylanase coding sequence having an altered codon usage from that of the template nucleic acid is produced.

10 140. The method of claim 131, wherein the method is iteratively repeated until a xylanase gene having higher or lower level of message expression or stability from that of the template nucleic acid is produced.

141. A method for modifying codons in a nucleic acid encoding a polypeptide with a xylanase activity to increase its expression in a host cell, the method
15 comprising the following steps:

(a) providing a nucleic acid encoding a polypeptide with a xylanase activity comprising a sequence as set forth in claim 1 or claim 24; and,

(b) identifying a non-preferred or a less preferred codon in the nucleic acid of step (a) and replacing it with a preferred or neutrally used codon encoding the same amino
20 acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to increase its expression in a host cell.

25 142. A method for modifying codons in a nucleic acid encoding a xylanase polypeptide, the method comprising the following steps:

(a) providing a nucleic acid encoding a polypeptide with a xylanase activity comprising a sequence as set forth in claim 1 or claim 24; and,

(b) identifying a codon in the nucleic acid of step (a) and replacing it with a
30 different codon encoding the same amino acid as the replaced codon, thereby modifying codons in a nucleic acid encoding a xylanase.

143. A method for modifying codons in a nucleic acid encoding a xylanase polypeptide to increase its expression in a host cell, the method comprising the following steps:

- 5 (a) providing a nucleic acid encoding a xylanase polypeptide comprising a sequence as set forth in claim 1 or claim 24; and,
- (b) identifying a non-preferred or a less preferred codon in the nucleic acid of step (a) and replacing it with a preferred or neutrally used codon encoding the same amino acid as the replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in the host cell and a non-preferred or less preferred codon is a codon
- 10 under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to increase its expression in a host cell.

144. A method for modifying a codon in a nucleic acid encoding a polypeptide having a xylanase activity to decrease its expression in a host cell, the method

15 comprising the following steps:

- (a) providing a nucleic acid encoding a xylanase polypeptide comprising a sequence as set forth in claim 1 or claim 24; and
- (b) identifying at least one preferred codon in the nucleic acid of step (a) and replacing it with a non-preferred or less preferred codon encoding the same amino acid as the
- 20 replaced codon, wherein a preferred codon is a codon over-represented in coding sequences in genes in a host cell and a non-preferred or less preferred codon is a codon under-represented in coding sequences in genes in the host cell, thereby modifying the nucleic acid to decrease its expression in a host cell.

25 145. The method of claim 144, wherein the host cell is a bacterial cell, a fungal cell, an insect cell, a yeast cell, a plant cell or a mammalian cell.

146. A method for producing a library of nucleic acids encoding a plurality of modified xylanase active sites or substrate binding sites, wherein the modified active sites or substrate binding sites are derived from a first nucleic acid comprising a sequence

30 encoding a first active site or a first substrate binding site the method comprising the following steps:

- (a) providing a first nucleic acid encoding a first active site or first substrate binding site, wherein the first nucleic acid sequence comprises a sequence that hybridizes

under stringent conditions to a sequence as set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:45, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:51, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:57, SEQ ID NO:59, SEQ ID NO:61, SEQ ID NO:63, SEQ ID NO:65, SEQ ID NO:67, SEQ ID NO:69, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:75, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:97, SEQ ID NO:99, SEQ ID NO:101, SEQ ID NO:103, SEQ ID NO:105, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NO:113, SEQ ID NO:115, SEQ ID NO:117, SEQ ID NO:119, SEQ ID NO:121, SEQ ID NO:123, SEQ ID NO:125, SEQ ID NO:127, SEQ ID NO:129, SEQ ID NO:131, SEQ ID NO:133, SEQ ID NO:135, SEQ ID NO:137, SEQ ID NO:139, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:157, SEQ ID NO:199, SEQ ID NO:161, SEQ ID NO:163, SEQ ID NO:165, SEQ ID NO:167, SEQ ID NO:169, SEQ ID NO:171, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:183, SEQ ID NO:185, SEQ ID NO:187, SEQ ID NO:189, SEQ ID NO:191, SEQ ID NO:193, SEQ ID NO:195, SEQ ID NO:197, SEQ ID NO:199, SEQ ID NO:201, SEQ ID NO:203, SEQ ID NO:205, SEQ ID NO:207, SEQ ID NO:209, SEQ ID NO:211, SEQ ID NO:213, SEQ ID NO:215, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:221, SEQ ID NO:223, SEQ ID NO:225, SEQ ID NO:227, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:233, SEQ ID NO:235, SEQ ID NO:237, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:243, SEQ ID NO:245, SEQ ID NO:247, SEQ ID NO:249, SEQ ID NO:251, SEQ ID NO:253, SEQ ID NO:255, SEQ ID NO:257, SEQ ID NO:259, SEQ ID NO:261, SEQ ID NO:263, SEQ ID NO:265, SEQ ID NO:267, SEQ ID NO:269, SEQ ID NO:271, SEQ ID NO:273, SEQ ID NO:275, SEQ ID NO:277, SEQ ID NO:279, SEQ ID NO:281, SEQ ID NO:283, SEQ ID NO:285, SEQ ID NO:287, SEQ ID NO:289, SEQ ID NO:291, SEQ ID NO:293, SEQ ID NO:295, SEQ ID NO:297, SEQ ID NO:299, SEQ ID NO:301, SEQ ID NO:303, SEQ ID NO:305, SEQ ID NO:307, SEQ ID NO:309, SEQ ID NO:311, SEQ ID NO:313, SEQ ID NO:315, SEQ ID NO:317, SEQ ID NO:319, SEQ ID NO:321, SEQ ID NO:323, SEQ ID NO:325, SEQ ID NO:327, SEQ ID NO:329, SEQ ID NO:331, SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:339, SEQ ID NO:341, SEQ ID NO:343, SEQ ID

NO:345, SEQ ID NO:347, SEQ ID NO:349, SEQ ID NO:351, SEQ ID NO:353, SEQ ID NO:355, SEQ ID NO:357, SEQ ID NO:359, SEQ ID NO:361, SEQ ID NO:363, SEQ ID NO:365, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:371, SEQ ID NO:373, SEQ ID NO:375, SEQ ID NO:377 or SEQ ID NO:379, or a subsequence thereof, and the nucleic acid
5 encodes a xylanase active site or a xylanase substrate binding site;

(b) providing a set of mutagenic oligonucleotides that encode naturally-occurring amino acid variants at a plurality of targeted codons in the first nucleic acid; and,

(c) using the set of mutagenic oligonucleotides to generate a set of active site-encoding or substrate binding site-encoding variant nucleic acids encoding a range of amino
10 acid variations at each amino acid codon that was mutagenized, thereby producing a library of nucleic acids encoding a plurality of modified xylanase active sites or substrate binding sites.

147. The method of claim 145, comprising mutagenizing the first nucleic
15 acid of step (a) by a method comprising an optimized directed evolution system, gene site-saturation mutagenesis (GSSM™), or a synthetic ligation reassembly (SLR).

148. The method of claim 145, comprising mutagenizing the first nucleic
acid of step (a) or variants by a method comprising error-prone PCR, shuffling,
20 oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, in vivo mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis, site-specific mutagenesis, gene reassembly, gene site saturated mutagenesis (GSSM™), synthetic ligation reassembly (SLR) and a combination thereof.

25 149. The method of claim 145, comprising mutagenizing the first nucleic acid of step (a) or variants by a method comprising recombination, recursive sequence recombination, phosphothioate-modified DNA mutagenesis, uracil-containing template mutagenesis, gapped duplex mutagenesis, point mismatch repair mutagenesis, repair-deficient host strain mutagenesis, chemical mutagenesis, radiogenic mutagenesis, deletion
30 mutagenesis, restriction-selection mutagenesis, restriction-purification mutagenesis, artificial gene synthesis, ensemble mutagenesis, chimeric nucleic acid multimer creation and a combination thereof.

150. A method for making a small molecule comprising the following steps:

(a) providing a plurality of biosynthetic enzymes capable of synthesizing or modifying a small molecule, wherein one of the enzymes comprises a xylanase enzyme encoded by a nucleic acid comprising a sequence as set forth in claim 1 or claim 24;

(b) providing a substrate for at least one of the enzymes of step (a); and

5 (c) reacting the substrate of step (b) with the enzymes under conditions that facilitate a plurality of biocatalytic reactions to generate a small molecule by a series of biocatalytic reactions.

10 151. A method for modifying a small molecule comprising the following steps:

(a) providing a xylanase enzyme, wherein the enzyme comprises a polypeptide as set forth in claim 64, or a polypeptide encoded by a nucleic acid comprising a nucleic acid sequence as set forth in claim 1 or claim 24;

(b) providing a small molecule; and

15 (c) reacting the enzyme of step (a) with the small molecule of step (b) under conditions that facilitate an enzymatic reaction catalyzed by the xylanase enzyme, thereby modifying a small molecule by a xylanase enzymatic reaction.

20 152. The method of claim 151, comprising a plurality of small molecule substrates for the enzyme of step (a), thereby generating a library of modified small molecules produced by at least one enzymatic reaction catalyzed by the xylanase enzyme.

25 153. The method of claim 151, further comprising a plurality of additional enzymes under conditions that facilitate a plurality of biocatalytic reactions by the enzymes to form a library of modified small molecules produced by the plurality of enzymatic reactions.

30 154. The method of claim 153, further comprising the step of testing the library to determine if a particular modified small molecule which exhibits a desired activity is present within the library.

155. The method of claim 154, wherein the step of testing the library further comprises the steps of systematically eliminating all but one of the biocatalytic reactions used to produce a portion of the plurality of the modified small molecules within the library by

testing the portion of the modified small molecule for the presence or absence of the particular modified small molecule with a desired activity, and identifying at least one specific biocatalytic reaction that produces the particular modified small molecule of desired activity.

5

156. A method for determining a functional fragment of a xylanase enzyme comprising the steps of:

(a) providing a xylanase enzyme, wherein the enzyme comprises a polypeptide as set forth in claim 64, or a polypeptide encoded by a nucleic acid as set forth in claim 1 or claim 24; and

(b) deleting a plurality of amino acid residues from the sequence of step (a) and testing the remaining subsequence for a xylanase activity, thereby determining a functional fragment of a xylanase enzyme.

157. The method of claim 156, wherein the xylanase activity is measured by providing a xylanase substrate and detecting a decrease in the amount of the substrate or an increase in the amount of a reaction product.

158. A method for whole cell engineering of new or modified phenotypes by using real-time metabolic flux analysis, the method comprising the following steps:

(a) making a modified cell by modifying the genetic composition of a cell, wherein the genetic composition is modified by addition to the cell of a nucleic acid comprising a sequence as set forth in claim 1 or claim 24;

(b) culturing the modified cell to generate a plurality of modified cells;

(c) measuring at least one metabolic parameter of the cell by monitoring the cell culture of step (b) in real time; and,

(d) analyzing the data of step (c) to determine if the measured parameter differs from a comparable measurement in an unmodified cell under similar conditions, thereby identifying an engineered phenotype in the cell using real-time metabolic flux analysis.

159. The method of claim 158, wherein the genetic composition of the cell is modified by a method comprising deletion of a sequence or modification of a sequence in the cell, or, knocking out the expression of a gene.

160. The method of claim 158, further comprising selecting a cell comprising a newly engineered phenotype.

5 161. The method of claim 160, further comprising culturing the selected cell, thereby generating a new cell strain comprising a newly engineered phenotype.

162. An isolated or recombinant signal sequence consisting of a sequence as set forth in residues 1 to 14, 1 to 15, 1 to 16, 1 to 17, 1 to 18, 1 to 19, 1 to 20, 1 to 21, 1 to 22, 1 to 23, 1 to 24, 1 to 25, 1 to 26, 1 to 27, 1 to 28, 1 to 28, 1 to 30, 1 to 31, 1 to 32, 1 to 33, 1 to 34, 1 to 35, 1 to 36, 1 to 37, 1 to 38, 1 to 40, 1 to 41, 1 to 42, 1 to 43 or 1 to 44, of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:46, SEQ ID NO:48, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:54, SEQ ID NO:56, SEQ ID NO:58, SEQ ID NO:60, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:66, SEQ ID NO:68, SEQ ID NO:70, SEQ ID NO:72, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:94, SEQ ID NO:96, SEQ ID NO:98, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:104, SEQ ID NO:106, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:112, SEQ ID NO:114, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:120, SEQ ID NO:122, SEQ ID NO:124, SEQ ID NO:126, SEQ ID NO:128, SEQ ID NO:130, SEQ ID NO:132; SEQ ID NO:134; SEQ ID NO:136; SEQ ID NO:138; SEQ ID NO:140; SEQ ID NO:142; SEQ ID NO:144; NO:146, SEQ ID NO:148, SEQ ID NO:150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:158, SEQ ID NO:160, SEQ ID NO:162, SEQ ID NO:164, SEQ ID NO:166, SEQ ID NO:168, SEQ ID NO:170, SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:180, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:186, SEQ ID NO:188, SEQ ID NO:190, SEQ ID NO:192, SEQ ID NO:194, SEQ ID NO:196, SEQ ID NO:198, SEQ ID NO:200, SEQ ID NO:202, SEQ ID NO:204, SEQ ID NO:206, SEQ ID NO:208, SEQ ID NO:210, SEQ ID NO:212, SEQ ID NO:214, SEQ ID NO:216, SEQ ID NO:218, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:224, SEQ ID NO:226, SEQ ID NO:228, SEQ ID NO:230, SEQ ID NO:232, SEQ ID NO:234, SEQ ID NO:236, SEQ ID NO:238, SEQ ID NO:240, SEQ ID

NO:242, SEQ ID NO:244, SEQ ID NO:246, SEQ ID NO:248, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NO:256, SEQ ID NO:258, SEQ ID NO:260, SEQ ID NO:262, SEQ ID NO:264, SEQ ID NO:266, SEQ ID NO:268, SEQ ID NO:270, SEQ ID NO:272, SEQ ID NO:274, SEQ ID NO:276, SEQ ID NO:278, SEQ ID NO:280, SEQ ID NO:282, SEQ ID NO:284, SEQ ID NO:286, SEQ ID NO:288, SEQ ID NO:290, SEQ ID NO:292, SEQ ID NO:294, SEQ ID NO:296, SEQ ID NO:298, SEQ ID NO:300, SEQ ID NO:302, SEQ ID NO:304, SEQ ID NO:306, SEQ ID NO:308, SEQ ID NO:310, SEQ ID NO:312, SEQ ID NO:314, SEQ ID NO:316, SEQ ID NO:318, SEQ ID NO:320, SEQ ID NO:322, SEQ ID NO:324, SEQ ID NO:326, SEQ ID NO:328, SEQ ID NO:330, SEQ ID NO:332, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:338, SEQ ID NO:340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:352, SEQ ID NO:354, SEQ ID NO:356, SEQ ID NO:358, SEQ ID NO:360, SEQ ID NO:362, SEQ ID NO:364, SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:372, SEQ ID NO:374, SEQ ID NO:376, SEQ ID NO:378 or SEQ ID NO:380; or,
consisting of a sequence as set forth in Table 4.

163. A chimeric polypeptide comprising at least a first domain comprising signal peptide (SP) having a sequence as set forth in claim 162, and at least a second domain comprising a heterologous polypeptide or peptide, wherein the heterologous polypeptide or peptide is not naturally associated with the signal peptide (SP).

164. The chimeric polypeptide of claim 163, wherein the heterologous polypeptide or peptide is not a xylanase.

165. The chimeric polypeptide of claim 163, wherein the heterologous polypeptide or peptide is amino terminal to, carboxy terminal to or on both ends of the signal peptide (SP) or a xylanase catalytic domain (CD).

166. An isolated or recombinant nucleic acid encoding a chimeric polypeptide, wherein the chimeric polypeptide comprises at least a first domain comprising signal peptide (SP) having a sequence as set forth in claim 162 and at least a second domain comprising a heterologous polypeptide or peptide, wherein the heterologous polypeptide or peptide is not naturally associated with the signal peptide (SP).

167. A method of increasing thermotolerance or thermostability of a xylanase polypeptide, the method comprising glycosylating a xylanase, wherein the polypeptide comprises at least thirty contiguous amino acids of a polypeptide as set forth in claim 60, or a polypeptide encoded by a nucleic acid as set forth in claim 1 or claim 24, thereby increasing the thermotolerance or thermostability of the xylanase.

168. A method for overexpressing a recombinant xylanase in a cell comprising expressing a vector comprising a nucleic acid sequence as set forth in claim 1 or claim 24, wherein overexpression is effected by use of a high activity promoter, a dicistronic vector or by gene amplification of the vector.

169. A method of making a transgenic plant comprising the following steps:
(a) introducing a heterologous nucleic acid sequence into the cell, wherein the heterologous nucleic sequence comprises a sequence as set forth in claim 1 or claim 24, thereby producing a transformed plant cell;
(b) producing a transgenic plant from the transformed cell.

170. The method as set forth in claim 169, wherein the step (a) further comprises introducing the heterologous nucleic acid sequence by electroporation or microinjection of plant cell protoplasts.

171. The method as set forth in claim 169, wherein the step (a) comprises introducing the heterologous nucleic acid sequence directly to plant tissue by DNA particle bombardment or by using an *Agrobacterium tumefaciens* host.

172. A method of expressing a heterologous nucleic acid sequence in a plant cell comprising the following steps:

(a) transforming the plant cell with a heterologous nucleic acid sequence operably linked to a promoter, wherein the heterologous nucleic sequence comprises a sequence as set forth in claim 1 or claim 24;
(b) growing the plant under conditions wherein the heterologous nucleic acids sequence is expressed in the plant cell.

173. A method for hydrolyzing, breaking up or disrupting a xylan-comprising composition comprising the following steps:

(a) providing a polypeptide having a xylanase activity as set forth in claim 64, or a polypeptide encoded by a nucleic acid as set forth in claim 1 or claim 24;

5 (b) providing a composition comprising a xylan; and

(c) contacting the polypeptide of step (a) with the composition of step (b) under conditions wherein the xylanase hydrolyzes, breaks up or disrupts the xylan-comprising composition.

10 174. The method as set forth in claim 173, wherein the composition comprises a plant cell, a bacterial cell, a yeast cell, an insect cell, or an animal cell.

175. A dough or a bread product comprising a polypeptide as set forth in claim 64.

15

176. A method of dough conditioning comprising contacting a dough or a bread product with at least one polypeptide as set forth in claim 64 under conditions sufficient for conditioning the dough.

20 177. A beverage comprising a polypeptide as set forth in claim 64.

178. A method of beverage production comprising administration of at least one polypeptide as set forth in claim 64 to a beverage or a beverage precursor under conditions sufficient for decreasing the viscosity of the beverage.

25

179. The method of claim 178, wherein the beverage or beverage precursor is a wort or a beer.

30 180. A food, a feed or a nutritional supplement comprising a polypeptide as set forth in claim 64.

181. A method for utilizing a xylanase as a nutritional supplement in an animal diet, the method comprising:

preparing a nutritional supplement containing a xylanase enzyme comprising at least thirty contiguous amino acids of a polypeptide as set forth in claim 64; and administering the nutritional supplement to an animal to increase utilization of a xylan contained in a feed or a food ingested by the animal.

5

182. The method of claim 181, wherein the animal is a human.

183. The method of claim 181, wherein the animal is a human.

10

184. The method of claim 181, wherein the animal is a ruminant or a monogastric animal.

15

185. The method of claim 181, wherein the xylanase enzyme is prepared by expression of a polynucleotide encoding the xylanase in an organism selected from the group consisting of a bacterium, a yeast, a plant, an insect, a fungus and an animal.

20

186. The method of claim 185, wherein the organism is selected from the group consisting of an *S. pombe*, *S. cerevisiae*, *Pichia pastoris*, *Pseudomonas* sp., *E. coli*, *Streptomyces* sp., *Bacillus* sp. and *Lactobacillus* sp.

25

187. An edible enzyme delivery matrix comprising a thermostable recombinant xylanase enzyme.

188. The edible enzyme delivery matrix of claim 187 comprising a polypeptide as set forth in claim 64.

30

189. A method for delivering a xylanase supplement to an animal, the method comprising:
preparing an edible enzyme delivery matrix in the form of pellets comprising a granulate edible carrier and a thermostable recombinant xylanase enzyme, wherein the pellets readily disperse the xylanase enzyme contained therein into aqueous media, and administering the edible enzyme delivery matrix to the animal.

190. The method of claim 189, wherein the recombinant xylanase enzyme comprises a polypeptide as set forth in claim 64.

5 191. The method of claim 189, wherein the granulate edible carrier comprises a carrier selected from the group consisting of a grain germ, a grain germ that is spent of oil, a hay, an alfalfa, a timothy, a soy hull, a sunflower seed meal and a wheat midd.

192. The method of claim 189, wherein the edible carrier comprises grain germ that is spent of oil.
10

193. The method of claim 189, wherein the xylanase enzyme is glycosylated to provide thermostability at pelletizing conditions.

194. The method of claim 189, wherein the delivery matrix is formed by pelletizing a mixture comprising a grain germ and a xylanase.
15

195. The method of claim 189, wherein the pelletizing conditions include application of steam.

20 196. The method of claim 189, wherein the pelletizing conditions comprise application of a temperature in excess of about 80°C for about 5 minutes and the enzyme retains a specific activity of at least 350 to about 900 units per milligram of enzyme.

25 197. An isolated or recombinant nucleic acid comprising a sequence encoding a polypeptide having a xylanase activity and a signal sequence, wherein the nucleic acid comprises a sequence as set forth in claim 1.

198. The isolated or recombinant nucleic acid of claim 197, wherein the signal sequence is derived from another xylanase or a non-xylanase enzyme.
30

199. An isolated or recombinant nucleic acid comprising a sequence encoding a polypeptide having a xylanase activity, wherein the sequence does not contain a signal sequence and the nucleic acid comprises a sequence as set forth in claim 1.

200. An isolated or recombinant nucleic acid comprising a sequence as set forth in SEQ ID NO: 189, wherein SEQ ID NO: 189 contains one or more of the following mutations: the nucleotides at positions 22 to 24 are TTC, the nucleotides at positions 31 to 33 are CAC, the nucleotides at positions 34 to 36 are TTG, the nucleotides at positions 49 to 51 are ATA, the nucleotides at positions 31 to 33 are CAT, the nucleotides at positions 67 to 69 are ACG, the nucleotides at positions 178 to 180 are CAC, the nucleotides at positions 190 to 192 are TGT, the nucleotides at positions 190 to 192 are GTA, the nucleotides at positions 190 to 192 are GTT, the nucleotides at positions 193 to 195 are GTG, the nucleotides at positions 202 to 204 are GCT, the nucleotides at positions 235 to 237 are CCA, or the nucleotides at positions 235 to 237 are CCC.

201. A method for making a nucleic acid comprising a sequence as set forth in claim 200, wherein the mutations in SEQ ID NO: 189 are obtained by gene site saturated mutagenesis (GSSM™).

15

202. An isolated or recombinant polypeptide comprising an amino acid sequence comprising SEQ ID NO: 190, wherein SEQ ID NO: 190 contains one or more of the following mutations: the aspartic acid at amino acid position 8 is phenylalanine, the glutamine at amino acid position 11 is histidine, the asparagine at amino acid position 12 is leucine, the glycine at amino acid position 17 is isoleucine, the threonine at amino acid position 23 is threonine encoded by a codon other than the wild type codon, the glycine at amino acid position 60 is histidine, the proline at amino acid position 64 is cysteine, the proline at amino acid position 64 is valine, the serine at amino acid position 65 is valine, the glycine at amino acid position 68 is isoleucine, the glycine at amino acid position 68 is alanine, or the valine at amino acid position 79 is proline.

203. A method for reducing lignin in a wood or wood product comprising contacting the wood or wood product with a polypeptide as set forth in claim 64.

204. A detergent composition comprising a polypeptide as set forth in claim 64.

205. A pharmaceutical composition comprising a polypeptide as set forth in claim 64.

206. A method for eliminating or protecting animals from a microorganism comprising a xylan comprising administering a polypeptide as set forth in claim 64.

5 207. The method of claim 206, wherein the microorganism is a bacterium.

208. The method of claim 205, wherein the bacterium is a salmonellae.

209. An isolated or recombinant nucleic acid comprising SEQ ID NO:189,
10 wherein SEQ ID NO:189 comprises one or more or all of the following sequence variations:
the nucleotides at positions 22 to 24 are TTC, the nucleotides at positions 22 to 24 are TTT,
the nucleotides at positions 31 to 33 are CAC, the nucleotides at positions 31 to 33 are CAT,
the nucleotides at positions 34 to 36 are TTG, the nucleotides at positions 34 to 36 are TTA,
the nucleotides at positions 34 to 36 are CTC, the nucleotides at positions 34 to 36 are CTT,
15 the nucleotides at positions 34 to 36 are CTA, the nucleotides at positions 34 to 36 are CTG,
the nucleotides at positions 49 to 51 are ATA, the nucleotides at positions 49 to 51 are ATT,
the nucleotides at positions 49 to 51 are ATC, the nucleotides at positions 178 to 180 are
CAC, the nucleotides at positions 178 to 180 are CAT, the nucleotides at positions 190 to 192
are TGT, the nucleotides at positions 190 to 192 are TGC, the nucleotides at positions 190 to
20 192 are GTA, the nucleotides at positions 190 to 192 are GTT, the nucleotides at positions
190 to 192 are GTC, the nucleotides at positions 190 to 192 are GTG, the nucleotides at
positions 193 to 195 are GTG, the nucleotides at positions 193 to 195 are GTC, the
nucleotides at positions 193 to 195 are GTA, the nucleotides at positions 193 to 195 are GTT,
the nucleotides at positions 202 to 204 are ATA, the nucleotides at positions 202 to 204 are
25 ATT, the nucleotides at positions 202 to 204 are ATC, the nucleotides at positions 202 to 204
are GCT, the nucleotides at positions 202 to 204 are GCG, the nucleotides at positions 202 to
204 are GCC, the nucleotides at positions 202 to 204 are GCA, the nucleotides at positions
235 to 237 are CCA, the nucleotides at positions 235 to 237 are CCC, or the nucleotides at
positions 235 to 237 are CCG.

30

210. An isolated or recombinant polypeptide comprising an amino acid
sequence comprising SEQ ID NO:190, wherein SEQ ID NO:190 comprises one or more or
all of the following sequence variations: the aspartic acid at amino acid position 8 is
phenylalanine, the glutamine at amino acid position 11 is histidine, the asparagine at amino

acid position 12 is leucine, the glycine at amino acid position 17 is isoleucine, the threonine at amino acid position 23 is threonine encoded by a codon other than the wild type codon, the glycine at amino acid position 60 is histidine, the proline at amino acid position 64 is cysteine, the proline at amino acid position 64 is valine, the serine at amino acid position 65 is valine, the glycine at amino acid position 68 is isoleucine, the glycine at amino acid position 68 is alanine, or the serine at amino acid position 79 is proline.

211. An isolated or recombinant nucleic acid comprising SEQ ID NO: 189, wherein SEQ ID NO:189 comprises one or more or all sequence variations set forth in Table 1 or Table 2.

212. An isolated or recombinant polypeptide encoded by the nucleic acid of claim 211.

213. An isolated or recombinant nucleic acid comprising SEQ ID NO:379, wherein SEQ ID NO:379 comprises one or more or all of the following sequence variations: the nucleotides at positions 22 to 24 are TTC, the nucleotides at positions 31 to 33 are CAC, the nucleotides at positions 49 to 51 are ATA, the nucleotides at positions 178 to 180 are CAC, the nucleotides at positions 193 to 195 are GTG, the nucleotides at positions 202 to 204 are GCT.

214. An isolated or recombinant polypeptide comprising SEQ ID NO:380, wherein SEQ ID NO:380 comprises one or more or all of the following sequence variations: D8F, Q11H, G17I, G60H, S65V and/or G68A.

215. The isolated or recombinant polypeptide of claim 210 or claim 214, wherein the polypeptide has a thermostable xylanase activity.

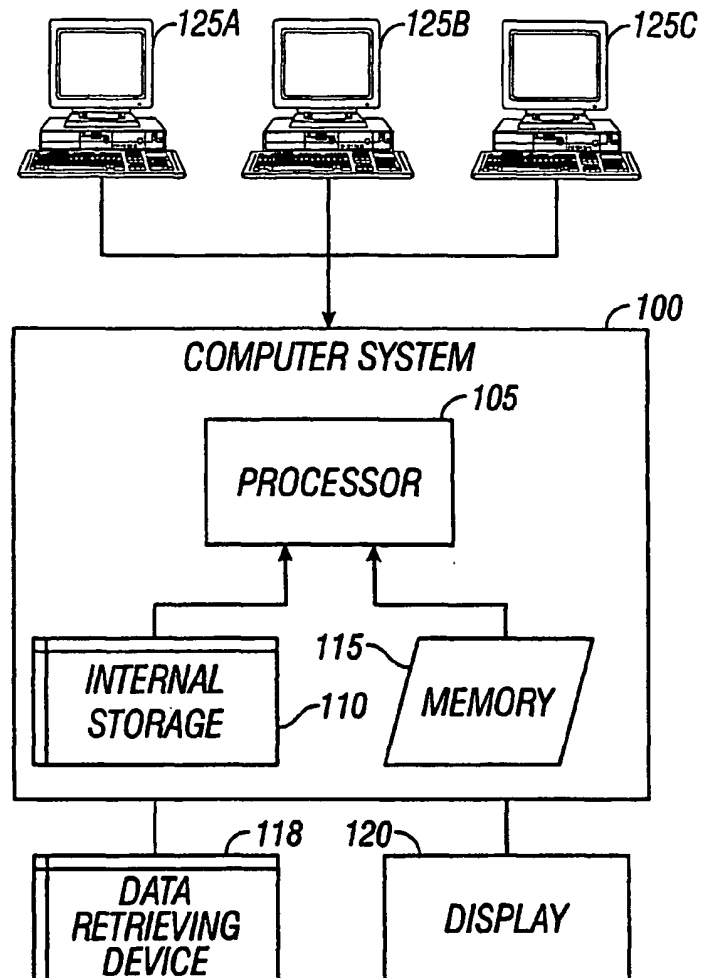


FIG. 1

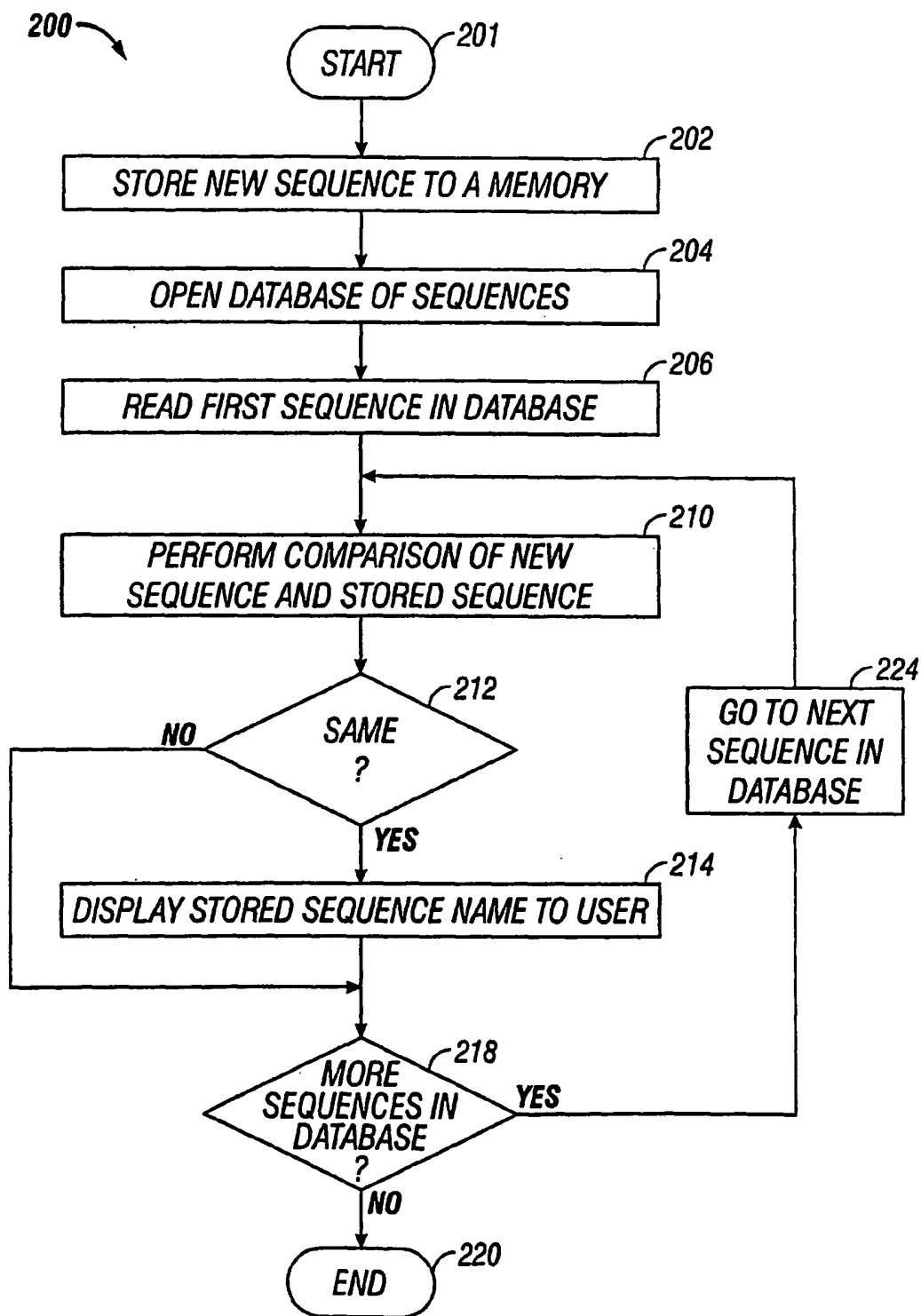


FIG. 2

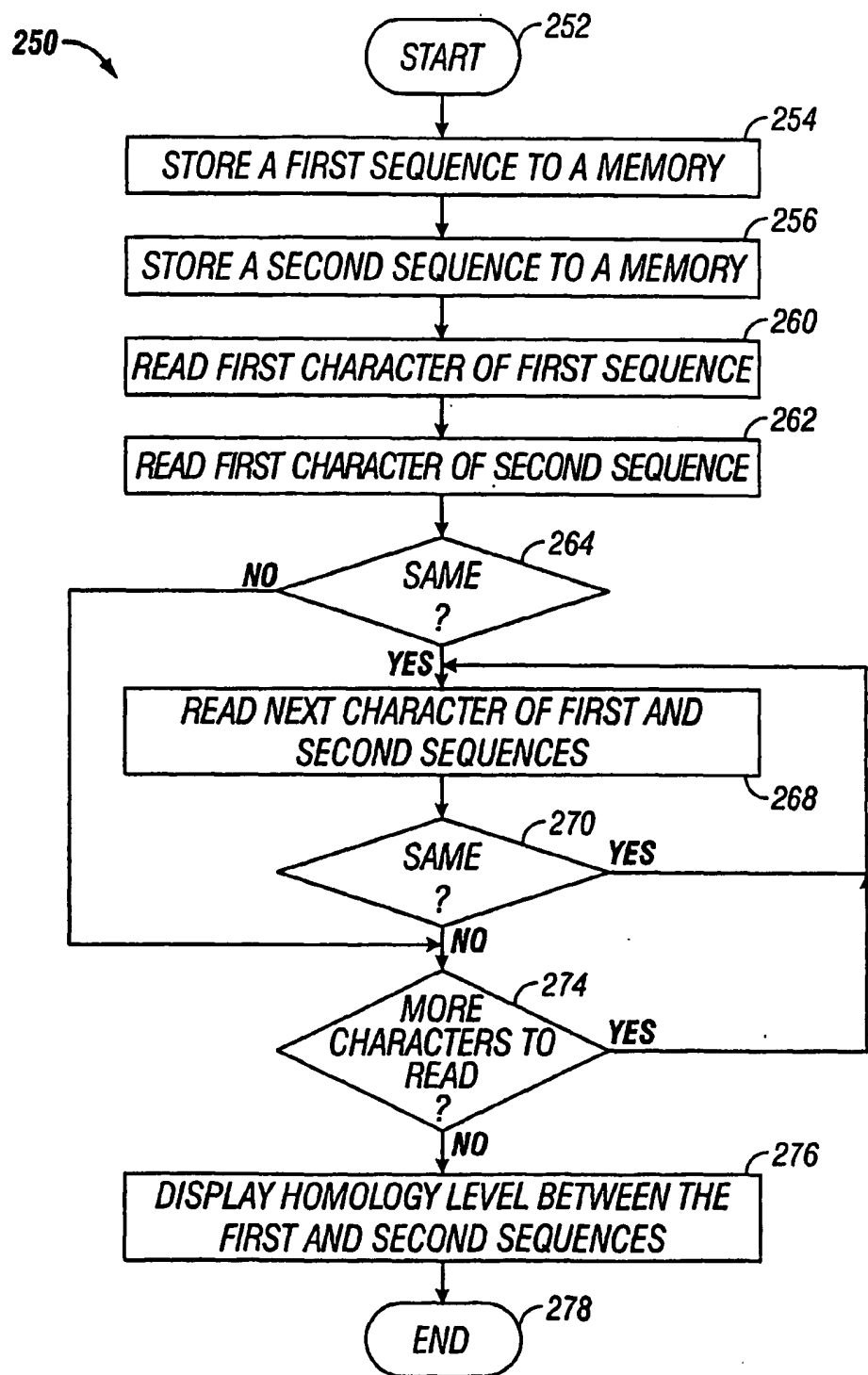


FIG. 3

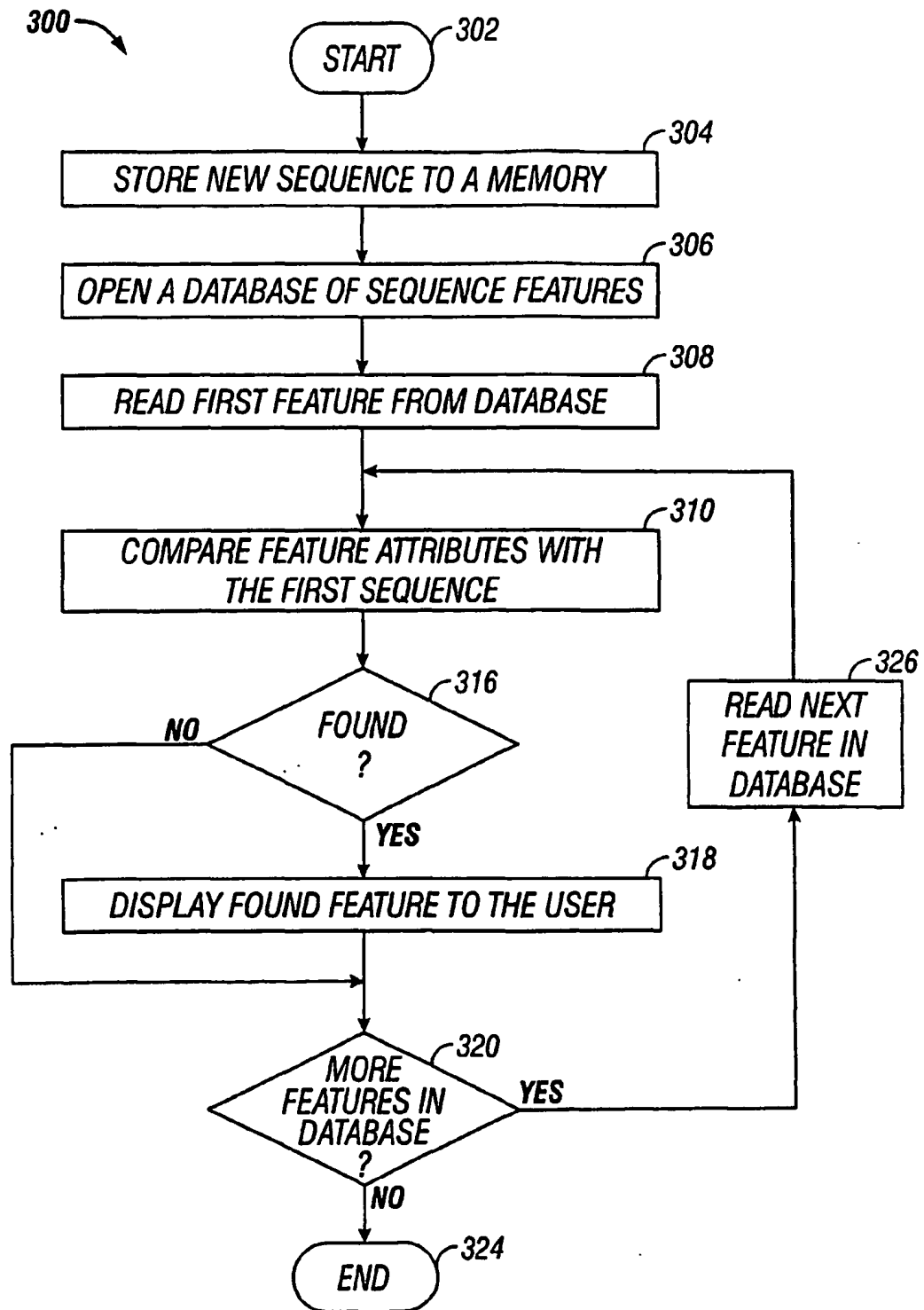


FIG. 4

**Figure 5: Thermal Tolerance of Wild-type Xylanase
(SEQ ID NOS:189 and 190)
vs. 8x Mutant**

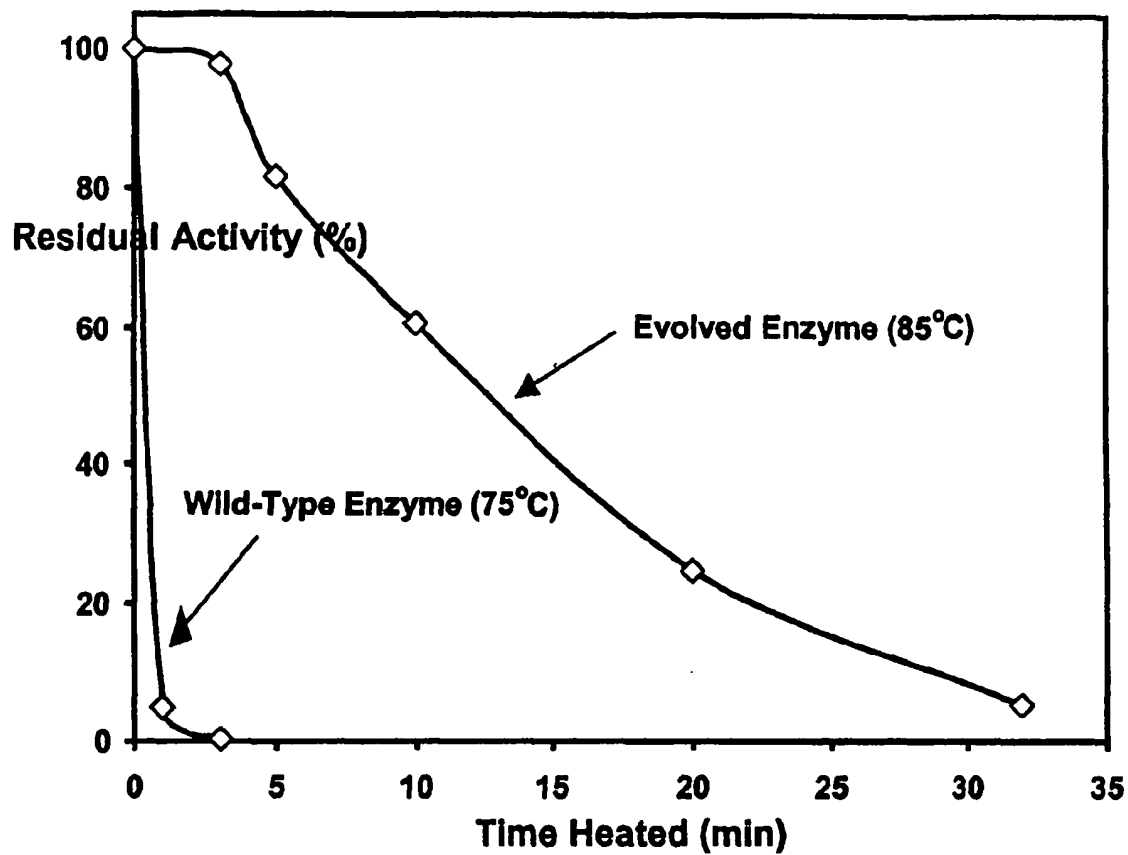


Figure 6

A

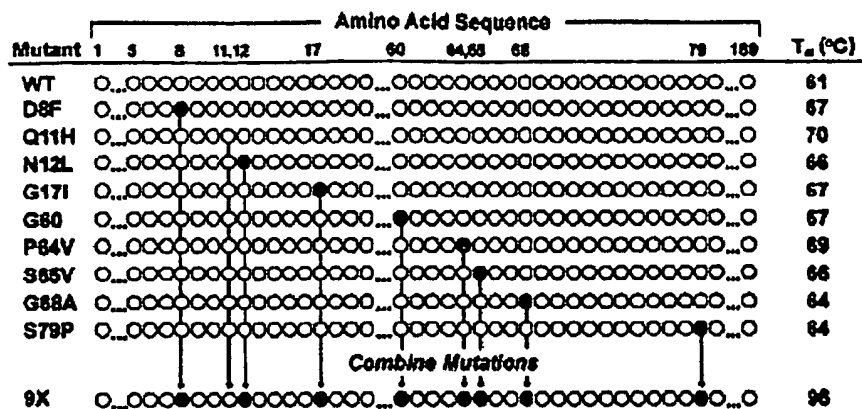
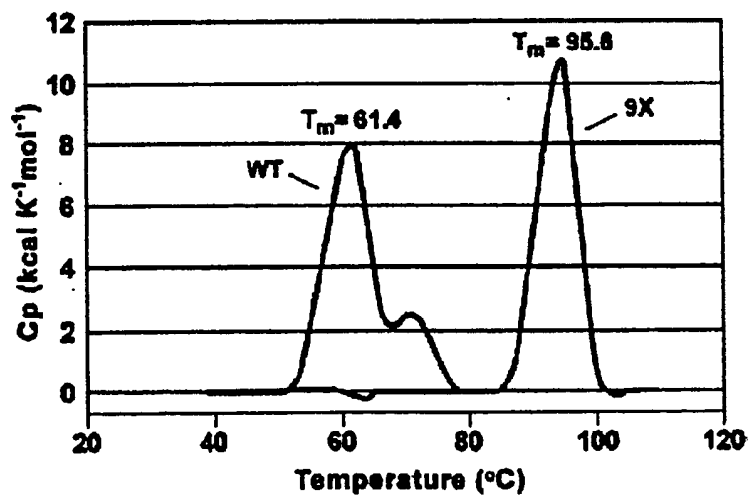
**B**

Figure 6

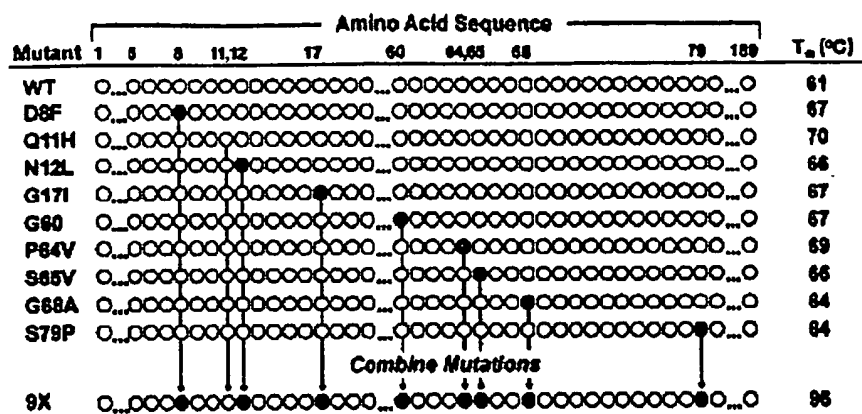
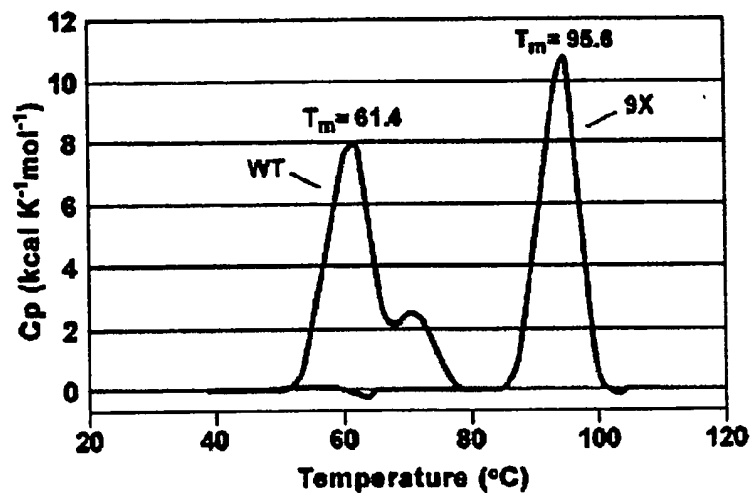
**B**

Figure 7

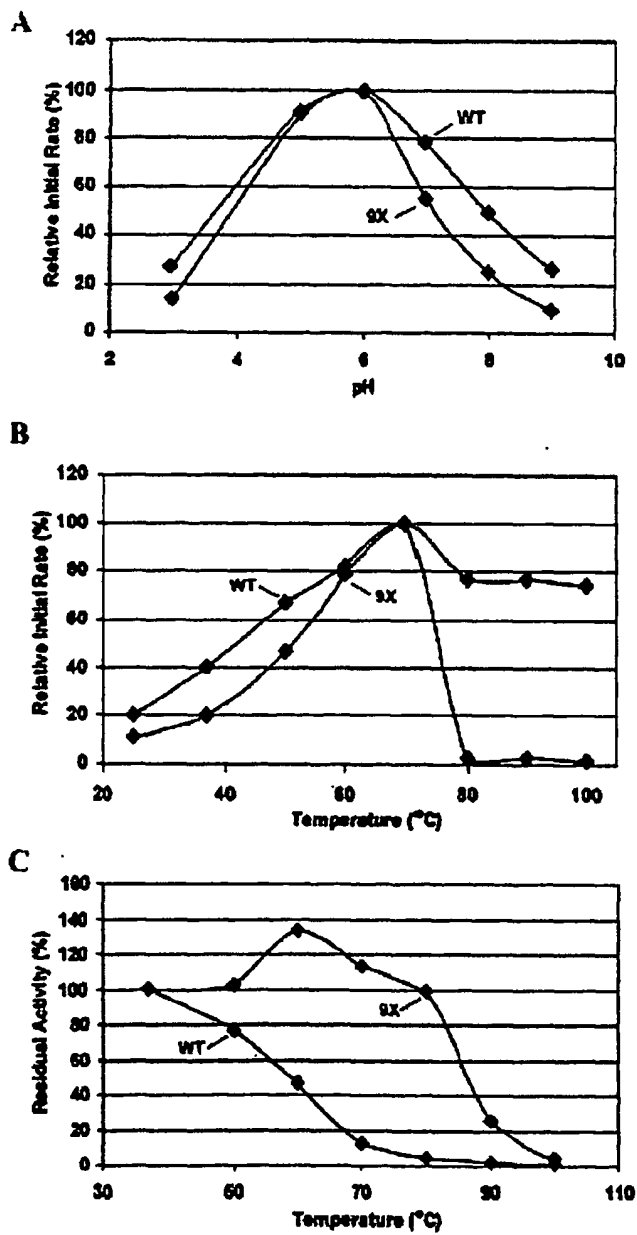


Figure 7

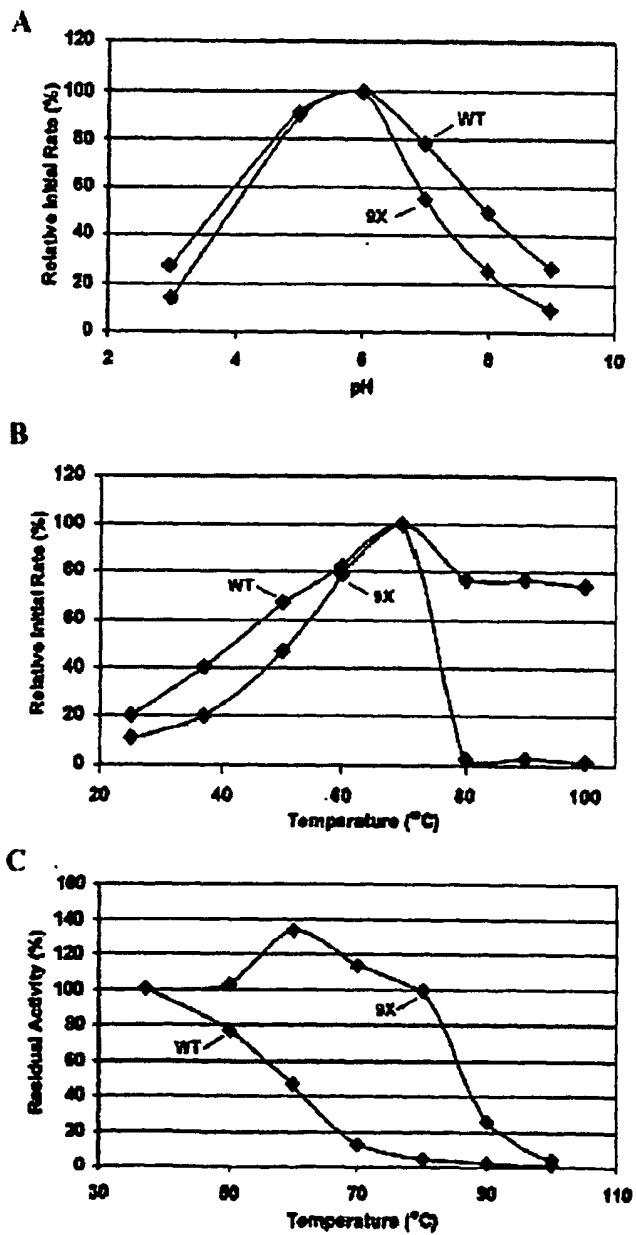


Figure 8

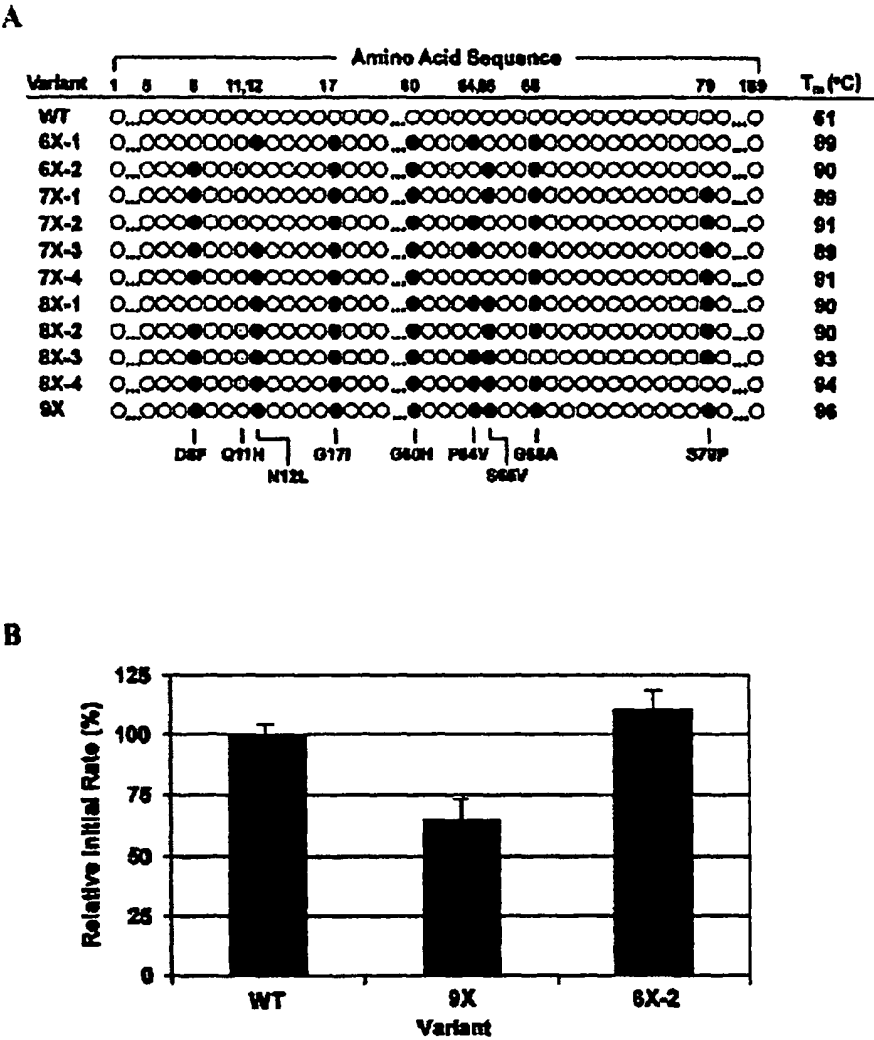


Figure 8

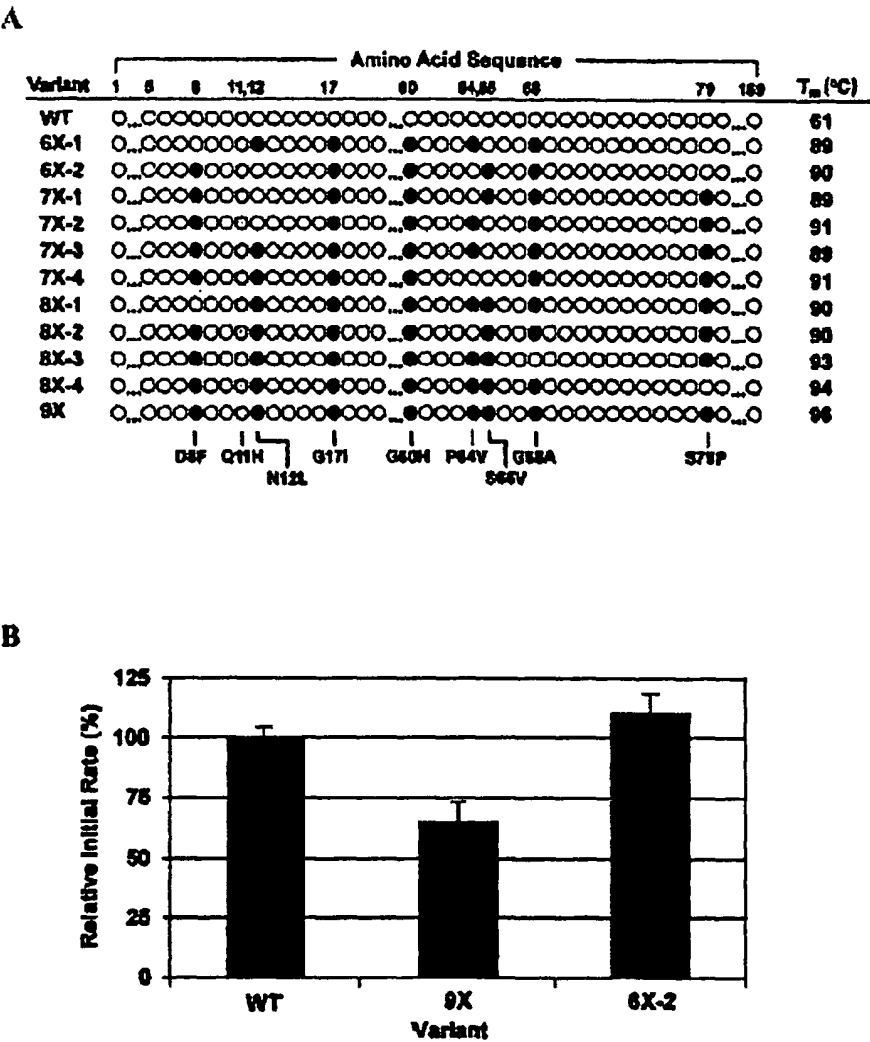


Figure 9

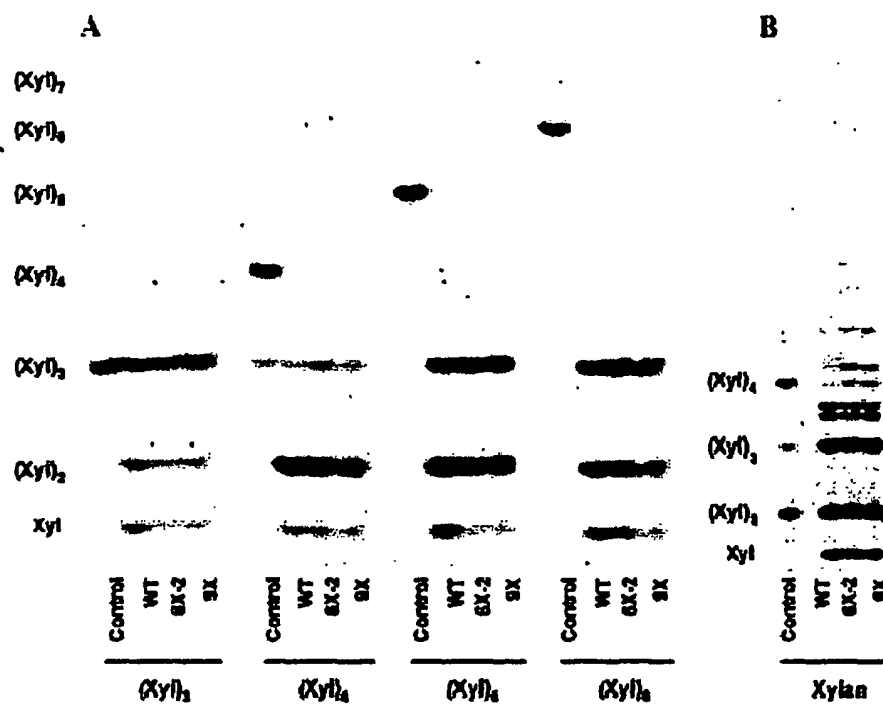


Figure 10

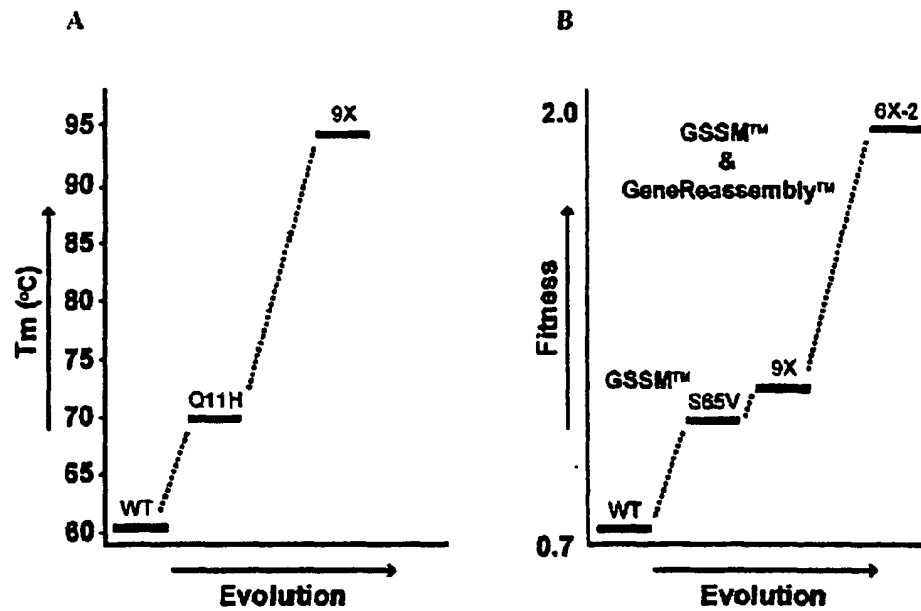


Figure 10

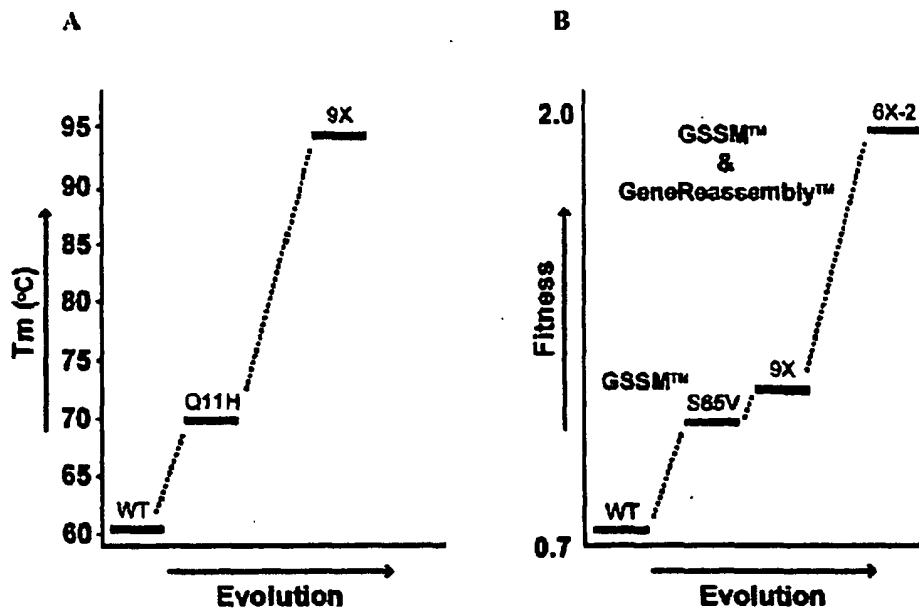
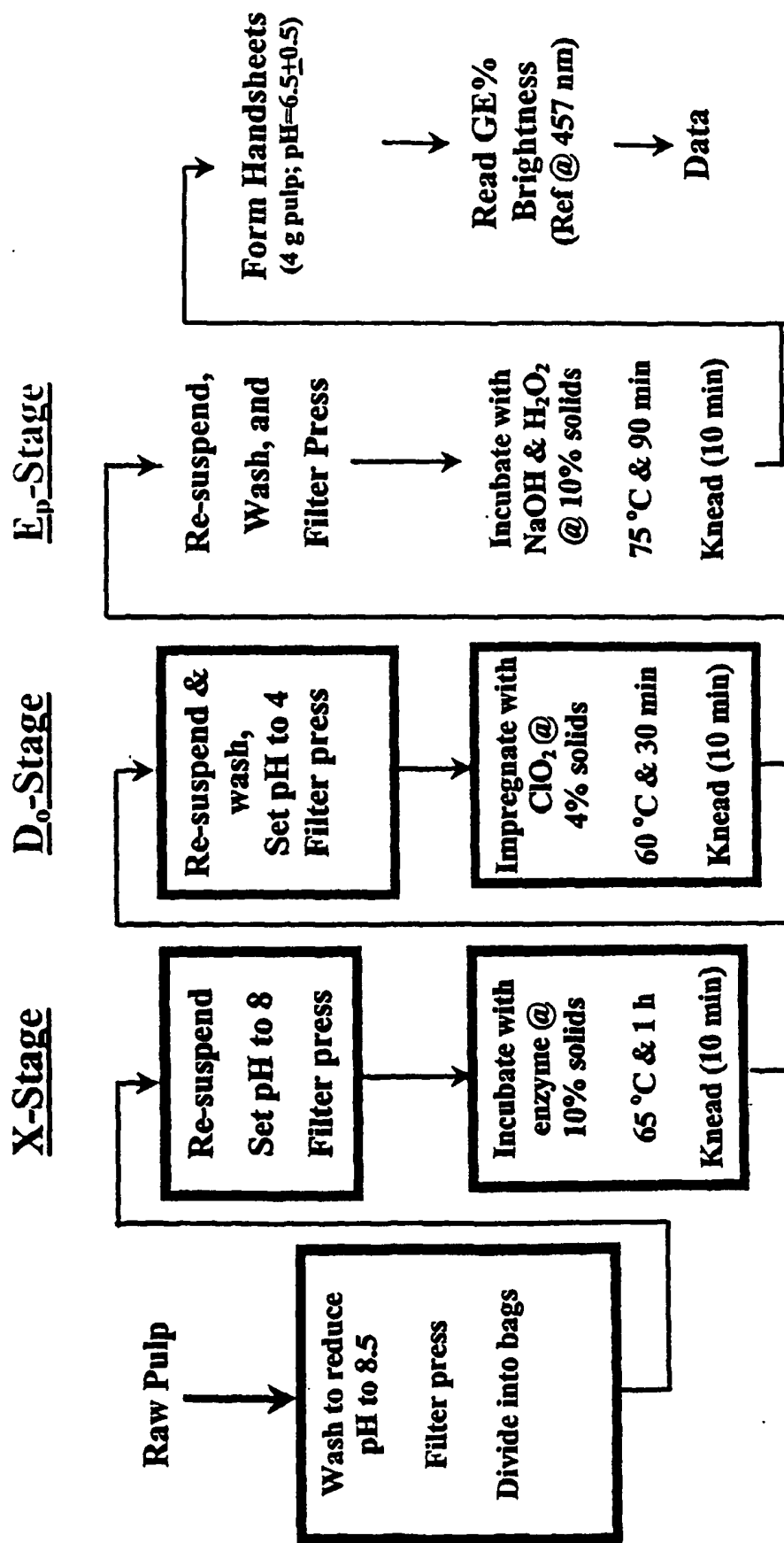


Fig. 11. Diversa Applications Lab Biobleaching Scheme



SEQUENCE LISTING

<110> Steer, Brian
 Callen, Walter
 Healey, Shaun
 Hazlewood, Geoff
 Wu, Di
 Blum, David
 Esteghlalian, Alireza

<120> XYLANASES, NUCLEIC ACIDS ENCODING THEM AND METHODS FOR MAKING AND USING THEM

<130> 09010-290001 and 09010-290W01

<140> not assigned

<141> 2003-06-16

<150> US 60/389,299

<151> 2002-06-14

<160> 380

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1128

<212> DNA

<213> Bacteria

<400> 1

atgaccgacc	acaacgcttc	cgaaccagc	ctgttcgaac	agtgcggcta	cagccgcgag	60
gccatccagg	cccgcctgga	gcgcaactgg	tatgagatgt	tcgaaggccc	ggacaagatt	120
tactgggaga	acgacgaagg	cctgggggtac	gtgatggaca	ccggcaacca	cgacgtgcgc	180
accgagggca	tgagctacgc	gatgatgatc	gccgtgcagt	acggccgcaa	ggacgtgttc	240
gacaagctgt	ggggttgggt	catgaaatac	atgttcatga	ccgagggcct	gcaccagggc	300
tacttcgcct	ggctctgtgga	ccccagcggc	gtaccgaacg	ccgacggctc	ggccccggac	360
ggcgaggaat	acttcgcgat	ggacctgttc	ctggcctccg	cgcgatgggg	cgacggcgaa	420
ggcgtgtacg	agtactcccg	ccacgcccgc	tcgatcctcc	acacctgcgt	gcaccagggc	480
gaggacgggtg	aaggctatcc	gatgtggaac	ccggagaacc	atctgatcaa	gttcattccc	540
gaaaccgaat	ggaccgaccc	gtcctaccat	ctgccgcact	tctacgaggt	gttcgccgag	600
cgcgccgacg	aggccgaccg	tccgttctgg	gcgcaggccg	ccaaggcgag	ccgcgagtac	660
ctgggtcaccg	cctgccaccc	gcagaccggc	atgaacccc	aataactcaa	ctatgatggc	720
acgcccgcacg	tcgacgagcg	cgaccactgg	catttctact	ccgacgccta	ccgcaccgcc	780
ggcaacatcg	ggctggactg	cctgtggaac	ggcgtcgtgc	cggaactgtg	cgatgcgaat	840
gcgcgtctgc	agcgtttctt	cctcgaacac	gaccgcacct	gcgtgtatgc	gatcgacggc	900
acgccgggtg	acgagaccgt	gctgcacccg	gtcggcttca	tcgccgccac	cgccgaaggc	960
tcgctcgccg	cgatgcactc	gcaggagccg	gacgcgctcg	acaacgcgat	ccgctgggtg	1020
cgctgtctgt	gggacacccc	gatccgcacc	ggcacgcgcc	gctactacga	caacttcctc	1080
tacgccttcg	cgttcctggc	gctggcgggg	gagtaccgca	cctgggtga		1128

<210> 2

<211> 375

<212> PRT

<213> Bacteria

<400> 2

Met	Thr	Asp	His	Asn	Ala	Ser	Glu	Thr	Ser	Leu	Phe	Glu	Gln	Cys	Gly
1				5					10					15	
Tyr	Ser	Arg	Glu	Ala	Ile	Gln	Ala	Arg	Leu	Glu	Arg	Asn	Trp	Tyr	Glu
		20						25					30		
Met	Phe	Glu	Gly	Pro	Asp	Lys	Ile	Tyr	Trp	Glu	Asn	Asp	Glu	Gly	Leu
		35					40				45				
Gly	Tyr	Val	Met	Asp	Thr	Gly	Asn	His	Asp	Val	Arg	Thr	Glu	Gly	Met
	50					55				60					
Ser	Tyr	Ala	Met	Met	Ile	Ala	Val	Gln	Tyr	Gly	Arg	Lys	Asp	Val	Phe
	65				70				75					80	
Asp	Lys	Leu	Trp	Gly	Trp	Val	Met	Lys	Tyr	Met	Phe	Met	Thr	Glu	Gly

Leu	His	Gln	Gly	Tyr	Phe	Ala	Trp	Ser	Val	Asp	Pro	Ser	Gly	Val	Pro
			100					105					110		
Asn	Ala	Asp	Gly	Pro	Ala	Pro	Asp	Gly	Glu	Glu	Tyr	Phe	Ala	Met	Asp
		115					120					125			
Leu	Phe	Leu	Ala	Ser	Ala	Arg	Trp	Gly	Asp	Gly	Glu	Gly	Val	Tyr	Glu
	130					135					140				
Tyr	Ser	Arg	His	Ala	Arg	Ser	Ile	Leu	His	Thr	Cys	Val	His	Gln	Gly
	145				150					155					160
Glu	Asp	Gly	Glu	Gly	Tyr	Pro	Met	Trp	Asn	Pro	Glu	Asn	His	Leu	Ile
				165					170					175	
Lys	Phe	Ile	Pro	Glu	Thr	Glu	Trp	Thr	Asp	Pro	Ser	Tyr	His	Leu	Pro
			180					185					190		
His	Phe	Tyr	Glu	Val	Phe	Ala	Glu	Arg	Ala	Asp	Glu	Ala	Asp	Arg	Pro
		195					200					205			
Phe	Trp	Ala	Gln	Ala	Ala	Lys	Ala	Ser	Arg	Glu	Tyr	Leu	Val	Thr	Ala
	210					215					220				
Cys	His	Pro	Gln	Thr	Gly	Met	Asn	Pro	Glu	Tyr	Ser	Asn	Tyr	Asp	Gly
	225				230					235					240
Thr	Pro	His	Val	Asp	Glu	Arg	Asp	His	Trp	His	Phe	Tyr	Ser	Asp	Ala
				245					250					255	
Tyr	Arg	Thr	Ala	Gly	Asn	Ile	Gly	Leu	Asp	Cys	Leu	Trp	Asn	Gly	Val
			260					265					270		
Val	Pro	Glu	Leu	Cys	Asp	Ala	Asn	Ala	Arg	Leu	Gln	Arg	Phe	Phe	Leu
		275					280					285			
Glu	His	Asp	Arg	Thr	Cys	Val	Tyr	Ala	Ile	Asp	Gly	Thr	Pro	Val	Asp
	290					295					300				
Glu	Thr	Val	Leu	His	Pro	Val	Gly	Phe	Ile	Ala	Ala	Thr	Ala	Glu	Gly
	305				310				315						320
Ser	Leu	Ala	Ala	Met	His	Ser	Gln	Glu	Pro	Asp	Ala	Leu	Asp	Asn	Ala
				325					330					335	
Ile	Arg	Trp	Val	Arg	Leu	Leu	Trp	Asp	Thr	Pro	Ile	Arg	Thr	Gly	Thr
			340					345					350		
Arg	Arg	Tyr	Tyr	Asp	Asn	Phe	Leu	Tyr	Ala	Phe	Ala	Phe	Leu	Ala	Leu
		355					360					365			
Ala	Gly	Glu	Tyr	Arg	Thr	Trp									
	370					375									

<210> 3
 <211> 2196
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 3

atgttcaaaag	aatgggggaaa	gaccgactcg	gaaattacca	caaaagtaaa	caccgcgtgg	60
aacaaactgt	ttgttaacgg	ggtggaatcc	ggagataatg	ccgaaagaat	ctatgtagag	120
actggaagcg	acatggcgta	catccatacc	tttgacagca	acgacgtgcg	ctccgaagga	180
atgtcctacg	gcatgatgat	gtgcgtacag	atgaacgatc	agacaagatt	taacaaactc	240
tggaatggg	caagaacctt	tatgtacaat	gaaacagacg	ccggcagtaa	ttccaggggc	300
tatttctcat	ggcagtgtag	tacaagcggc	tcaaaaatgg	ataagggccc	cgctcctgac	360
ggcgaggaat	actttattac	ggcgtgtgtg	ttcgcgcacg	cccgtggggg	gagcgcgtcc	420
ggtactacaa	acataaacia	ttacgcgcag	caagcaaggc	agattatcta	tgacttaacg	480
cgccgcaaac	cggggaacgg	agatccttac	ggcgcgcctt	caatgtttta	tgtagacaac	540
tatatgggta	gatttgcacc	acttggaaat	tccgccacct	ttacagacct	ctcataccat	600
ttaccggcat	tctatgatgt	ttgggcgctg	gaattacagg	cggactatga	taatagtaaa	660
ctctacggta	tctgggctga	taaggctgac	ttgaaaaaag	acattgatta	ctttaaacia	720
gcggcgacca	caagccgttc	attcctttga	aaaacgacaa	acggtacaac	cggacttgga	780
ccggattatg	ccggctttga	cggacgcctt	aaaaatgaag	gggatcacaa	gtatttcgag	840
tatgacgcgt	ggcgtatcgc	gatgaacata	ggtatggact	acgcgtgggt	cgcgaaagat	900
tcctggcaga	agacatttgc	cgacagaatt	caggcgttct	ttgtcagcaa	gggagtcact	960
tcttacggaa	accgctggac	attggacggg	actcaaaggg	gagcggatca	ctcgccgggt	1020
ctgtcgggt	gtaacgcggg	cgctctcttc	gcggcgacaa	acgcgaacgc	gtggaaattt	1080
atcggaagact	tctggaacat	cagcatgacg	aaaggcaaat	accgttacta	tgacggatgt	1140
ctgtatatga	tgagcatgct	gcacttaagc	ggcaacttta	aggcgtatct	ttctacaaat	1200
accacgcccg	ccaacagttc	cagcattacc	ccgacaaccg	cgtctttcga	caagaagaca	1260
agcgcaacaag	ccgacattgc	cgtaacagtg	acgcttaacg	ggaatacatt	ctcaagtatc	1320

acaaacaacg	gtacagccct	tacaagcggc	acagactact	cagtgagtgg	aacaaagtat	1380
acgataaaga	aagaatacct	tgcaaaacag	cctgtaggaa	caacgaagct	cgcatccaac	1440
ttcagtgccg	gaggaactcc	ggaacttaca	gttactataa	cggaacggg	cagctccagc	1500
atcagcccga	caaccgcgac	attcgacaaa	aagaccggag	cgcaagccga	catcgccgta	1560
accatgacgc	ttaatgggaa	tactttgtcg	aacatcaaaa	acggttctgc	acaacttaca	1620
agcggaactg	actactcaac	gagcggcagt	acggtaacga	ttaaaaaaga	atacctggca	1680
aagcaggcta	acggcacagt	aacgcttacc	ttcacattca	gcgcaggcgc	ggcccaaact	1740
attgacatca	cggtaaaaga	tacaaccggc	ggagcggcgg	gaataaaaata	caacttcgca	1800
actgacaacc	tgcccaacgg	gtacccgaag	tacagttcaa	gtgatataatc	cgcgacaata	1860
accggaggag	ctttggtaat	aacccaaaacc	ggaaataatt	cgtcccccga	gattacattg	1920
cccttttagtg	taacaggtaa	cctttccggt	tatacaggca	taaagataaa	tgtaaaggga	1980
gtatccggag	atttcactta	taaagtattg	aatgccgcaa	taggttctac	aaatctcggc	2040
agcgtaaata	acgccccaat	accaaaccggc	tcatttggag	acgtaacaat	accaataacc	2100
ggcggtagaa	acaccggaga	tttagatata	tcgttctggc	tcaataacac	aaatgcttac	2160
gttattgaga	ttaagagcat	agagctggta	aaatga			2196

<210> 4
 <211> 711
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 4
 Met Phe Lys Glu Trp Gly Lys Thr Asp Ser Glu Ile Thr Thr Lys Val
 1 5 10 15
 Asn Thr Ala Trp Asn Lys Leu Phe Val Asn Gly Val Glu Ser Gly Asp
 20 25 30
 Asn Ala Glu Arg Ile Tyr Val Glu Thr Gly Ser Asp Met Ala Tyr Ile
 35 40 45
 His Thr Phe Asp Ser Asn Asp Val Arg Ser Glu Gly Met Ser Tyr Gly
 50 55 60
 Met Met Met Cys Val Gln Met Asn Asp Gln Thr Arg Phe Asn Lys Leu
 65 70 75 80
 Trp Lys Trp Ala Arg Thr Tyr Met Tyr Asn Glu Thr Asp Ala Gly Ser
 85 90 95
 Asn Ser Arg Gly Tyr Phe Ser Trp Gln Cys Ser Thr Ser Gly Ser Lys
 100 105 110
 Met Asp Lys Gly Pro Ala Pro Asp Gly Glu Glu Tyr Phe Ile Thr Ala
 115 120 125
 Leu Leu Phe Ala His Ala Arg Trp Gly Ser Ala Ser Gly Thr Thr Asn
 130 135 140
 Ile Asn Asn Tyr Ala Gln Gln Ala Arg Gln Ile Ile Tyr Asp Leu Thr
 145 150 155 160
 Arg Arg Lys Pro Gly Asn Gly Asp Pro Tyr Gly Glu Pro Ser Met Phe
 165 170 175
 Asn Val Asp Asn Tyr Met Val Arg Phe Ala Thr Leu Gly Asn Ser Ala
 180 185 190
 Thr Phe Thr Asp Pro Ser Tyr His Leu Pro Ala Phe Tyr Asp Val Trp
 195 200 205
 Ala Leu Glu Leu Gln Ala Asp Tyr Asp Asn Ser Lys Leu Tyr Gly Ile
 210 215 220
 Trp Ala Asp Lys Ala Asp Leu Lys Lys Asp Ile Asp Tyr Phe Lys Gln
 225 230 235 240
 Ala Ala Thr Thr Ser Arg Ser Phe Phe Ala Lys Thr Thr Asn Gly Thr
 245 250 255
 Thr Gly Leu Gly Pro Asp Tyr Ala Gly Phe Asp Gly Thr Pro Lys Asn
 260 265 270
 Glu Gly Asp His Lys Tyr Phe Glu Tyr Asp Ala Trp Arg Ile Ala Met
 275 280 285
 Asn Ile Gly Met Asp Tyr Ala Trp Phe Ala Lys Asp Ser Trp Gln Lys
 290 295 300
 Thr Phe Ala Asp Arg Ile Gln Ala Phe Phe Val Ser Lys Gly Val Thr
 305 310 315 320
 Ser Tyr Gly Asn Arg Trp Thr Leu Asp Gly Thr Gln Arg Gly Ala Asp
 325 330 335
 His Ser Pro Gly Leu Val Gly Cys Asn Ala Val Ala Ser Leu Ala Ala
 340 345 350

Thr Asn Ala Asn Ala Trp Lys Phe Ile Glu Asp Phe Trp Asn Ile Ser
 355 360 365
 Met Thr Lys Gly Lys Tyr Arg Tyr Tyr Asp Gly Cys Leu Tyr Met Met
 370 375 380
 Ser Met Leu His Leu Ser Gly Asn Phe Lys Ala Tyr Leu Ser Thr Asn
 385 390 395 400
 Thr Thr Pro Ala Asn Ser Ser Ser Ile Thr Pro Thr Thr Ala Ser Phe
 405 410 415
 Asp Lys Lys Thr Ser Ala Gln Ala Asp Ile Ala Val Thr Val Thr Leu
 420 425 430
 Asn Gly Asn Thr Phe Ser Ser Ile Thr Asn Asn Gly Thr Ala Leu Thr
 435 440 445
 Ser Gly Thr Asp Tyr Ser Val Ser Gly Thr Lys Tyr Thr Ile Lys Lys
 450 455 460
 Glu Tyr Leu Ala Lys Gln Pro Val Gly Thr Thr Lys Leu Ala Phe Asn
 465 470 475 480
 Phe Ser Ala Gly Gly Thr Pro Glu Leu Thr Val Thr Ile Thr Asp Thr
 485 490 495
 Gly Ser Ser Ser Ile Ser Pro Thr Thr Ala Thr Phe Asp Lys Lys Thr
 500 505 510
 Gly Ala Gln Ala Asp Ile Ala Val Thr Met Thr Leu Asn Gly Asn Thr
 515 520 525
 Leu Ser Asn Ile Lys Asn Gly Ser Ala Gln Leu Thr Ser Gly Thr Asp
 530 535 540
 Tyr Ser Thr Ser Gly Ser Thr Val Thr Ile Lys Lys Glu Tyr Leu Ala
 545 550 555 560
 Lys Gln Ala Asn Gly Thr Val Thr Leu Thr Phe Thr Phe Ser Ala Gly
 565 570 575
 Ala Ala Gln Thr Ile Asp Ile Thr Val Lys Asp Thr Thr Gly Gly Ala
 580 585 590
 Ala Gly Ile Lys Tyr Asn Phe Ala Thr Asp Asn Leu Pro Asn Gly Tyr
 595 600 605
 Pro Lys Tyr Ser Ser Ser Asp Ile Ser Ala Thr Ile Thr Gly Gly Ala
 610 615 620
 Leu Val Ile Thr Lys Thr Gly Asn Asn Ser Ser Pro Lys Ile Thr Leu
 625 630 635 640
 Pro Phe Ser Val Thr Gly Asn Leu Ser Gly Tyr Thr Gly Ile Lys Ile
 645 650 655
 Asn Val Lys Gly Val Ser Gly Asp Phe Thr Tyr Lys Val Leu Asn Ala
 660 665 670
 Ala Ile Gly Ser Thr Asn Leu Gly Ser Val Asn Asn Ala Pro Ile Pro
 675 680 685
 Asn Gly Ser Phe Gly Asp Val Thr Ile Pro Ile Thr Gly Gly Thr Asn
 690 695 700
 Thr Gly Asp Leu Asp Ile Ser
 705 710

<210> 5

<211> 2106

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 5

atgcaaaacc	tatttaagcg	tgtgtttttc	catcttctct	tgcttgccct	gctggcaggc	60
tgtgctggcc	cttctcccgt	aacaccggag	ccgaccgaaa	tgccgaccga	ggtccctaca	120
ccaacgccta	gtcttgccgc	ctacgagagc	ggcgagtatc	gcaacctgtt	cgccgaggcg	180
cttggaacaat	cggatgccga	aattcaggcc	aaaatcgatg	ccgctttcca	acaacttttc	240
tacggcgacg	atgtttctga	gcgcgtctat	taccgggttg	gcagcgacat	gggctatatg	300
ctcgacaccg	gcaacgacga	tgtgcgctcc	gagggcatgt	cctacggcat	gatgattgcc	360
gtccagatga	acaagaagga	agaattcgac	cgcattctgga	agtggaccaa	aacctacatg	420
taccagaccg	aaggtggtta	caaagggttat	tttgccctggc	acgctaaaac	ggacggcacc	480
caactggccg	ccaacccggc	ctctgacggt	gaagtctggt	ttgtgatggc	gctcttcttt	540
gccgatgcgc	gttggggcag	cggcgaagga	atttataact	accgcgcccc	agcccaggaa	600
attctcgaatg	tggccttgaa	cgccaaagaa	ttgggcggca	acctggcgac	caacctgttc	660
gaccgcggaga	ccaaacaggt	cgtttttgtg	ccgcagtggg	gcaataactc	gaaatttacc	720
gacgcttcgt	accacatgcc	ccatttctac	gagtttgagg	cgcgttgggc	cgataaaaat	780

```

aacgactttt gggccgaagc cgctaccggt agccgcgagt tcctgcctac tgccgttcac      840
cccgaaaccg gcctggcccc taactattcc tacttcgatg gccgccctta caatgacgag      900
tatcacggcc agttccgcta cgacgctttc cgcgtgggcg cgaacatcgg catggattat      960
gtctggtttc acccctctga atggtatcgg gaacaagcca accgccaatt atctttcttc     1020
gcatcccagg gcatcgatga ttatgttgcc gaatattccc tggatggaaa accgctggcc     1080
gggcatcgcg ctacgggggtt gattgccacc aatgctgtcc tggcctacgc cgcagacccc     1140
gaaattggtc aacccttcgt ccaggccctg tgggatgcag agcctccgac tggcaggtat     1200
cgctactatg acggcctgct ctacatgatg ggcctgctgc aagccagcgg caacttcctg     1260
atttacgagc cgggtattac gcctcgcgct gagttgccgc ccccgccgcc tcgcgccatc     1320
gagggccgct tcgcgcccat taccgggcgg gccttgcttc tgattggccc gaatgcggat     1380
ggcgtcaacg cttacttcga caaactgggt acagcgccgg gcggcgtgaa tgtcgaacta     1440
tcgctcaaat cgcctgattt ggaagcgctc gacgccctgg cgaggaaata tcccaacagc     1500
acgctttcgg tcgggttgct gctggatggc ccggtaacag aggcggatgc gcgggtggga     1560
gaattgctcg acgcgttggc tgtttatccg cggccggtct tcctgcgcac cgggcccggaa     1620
tttgatttgg cggcgagcgg ccaggggcgg gaggaatatg tcgcggcctg gaaaacgctc     1680
cataacgaga ttacggcgcg gggtagttcg aatatcgccc tgggtgtggca tagcgccgca     1740
gcctgcgagt cgcccttttg cgtcatcccg ctcgaagcgt ggtatcccgg tgatgagttt     1800
gtggattggg tggccgtttc gcgcactgcg cagtctgcgg attgcgaggg gcagtcggtt     1860
gagggcgctc tgcagtttgc gcgtgagcga tacaaccggg ttgtgttggg tgcattcgca     1920
gcagaggaca tcttcgagtt cgtttacgcc aacaacgacg tgattcgcg cctgctgtat     1980
ctgaacaccg agccgggcct gttcgacacc cccgaatttt tgagcggctg gaaggccgaa     2040
atcggtcagc agttctggct gcgcggcggc ccggcgcttt tttcgacact cggattggat     2100
gagtaa

```

<210> 6

<211> 701

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(47)

<400> 6

```

Met Gln Asn Leu Phe Lys Arg Val Phe Phe His Leu Leu Leu Leu Ala
 1      5      10      15
Leu Leu Ala Gly Cys Ala Gly Pro Ser Pro Val Thr Pro Glu Pro Thr
      20      25      30
Glu Met Pro Thr Gln Val Pro Thr Pro Ser Leu Gly Ala Tyr
      35      40      45
Glu Ser Gly Glu Tyr Arg Asn Leu Phe Ala Glu Ala Leu Gly Lys Ser
      50      55      60
Asp Ala Glu Ile Gln Ala Lys Ile Asp Ala Ala Phe Gln Gln Leu Phe
      65      70      75      80
Tyr Gly Asp Asp Val Ser Glu Arg Val Tyr Tyr Pro Val Gly Ser Asp
      85      90      95
Met Gly Tyr Met Leu Asp Thr Gly Asn Asp Asp Val Arg Ser Glu Gly
      100      105      110
Met Ser Tyr Gly Met Met Ile Ala Val Gln Met Asn Lys Lys Glu Glu
      115      120      125
Phe Asp Arg Ile Trp Lys Trp Thr Lys Thr Tyr Met Tyr Gln Thr Glu
      130      135      140
Gly Gly Tyr Lys Gly Tyr Phe Ala Trp His Ala Lys Thr Asp Gly Thr
      145      150      155      160
Gln Leu Ala Ala Asn Pro Ala Ser Asp Gly Glu Val Trp Phe Val Met
      165      170      175
Ala Leu Phe Phe Ala Asp Ala Arg Trp Gly Ser Gly Glu Gly Ile Tyr
      180      185      190
Asn Tyr Arg Ala Gln Ala Gln Glu Ile Leu Asp Val Ala Leu Asn Ala
      195      200      205
Lys Glu Leu Gly Gly Asn Leu Ala Thr Asn Leu Phe Asp Pro Glu Thr
      210      215      220
Lys Gln Val Val Phe Val Pro Gln Leu Gly Asn Ser Lys Phe Thr
      225      230      235      240
Asp Ala Ser Tyr His Met Pro His Phe Tyr Glu Leu Trp Ala Arg Trp
      245      250      255
Ala Asp Lys Asn Asn Asp Phe Trp Ala Glu Ala Ala Thr Val Ser Arg

```


260 265 270
 Glu Phe Leu Pro Thr Ala Val His Pro Glu Thr Gly Leu Ala Pro Asn
 Tyr Ser 275 Phe Asp Gly Arg 280 Tyr Asn Asp Glu 285 Tyr His Gly Gln
 Phe Arg Tyr Asp Ala Phe Arg Val Gly Ala Asn Ile Gly Met Asp Tyr
 305 310 315 320
 Val Trp Phe His Pro Ser Glu Trp Tyr Arg Glu Gln Ala Asn Arg Gln
 325 330 335
 Leu Ser Phe Phe Ala Ser Gln Gly Ile Asp Asp Tyr Val Ala Glu Tyr
 340 345 350
 Ser Leu Asp Gly Lys Pro Leu Ala Gly His Arg Ala Thr Gly Leu Ile
 355 360 365
 Ala Thr Asn Ala Val Leu Ala Tyr Ala Ala Asp Pro Glu Ile Gly Gln
 370 375 380
 Pro Phe Val Gln Ala Leu Trp Asp Ala Glu Pro Pro Thr Gly Arg Tyr
 385 390 395 400
 Arg Tyr Tyr Asp Gly Leu Leu Tyr Met Met Gly Leu Leu Gln Ala Ser
 405 410 415
 Gly Asn Phe Arg Ile Tyr Glu Pro Gly Ile Thr Pro Arg Ala Glu Leu
 420 425 430
 Pro Pro Pro Pro Pro Arg Ala Ile Glu Gly Arg Phe Ala Pro Ile Thr
 435 440 445
 Gly Arg Ala Leu Leu Leu Ile Gly Pro Asn Ala Asp Gly Val Asn Ala
 450 455 460
 Tyr Phe Asp Lys Leu Val Thr Ala Pro Gly Gly Val Asn Val Glu Leu
 465 470 475 480
 Ser Leu Lys Ser Pro Asp Leu Glu Ala Leu Asp Ala Leu Ala Arg Lys
 485 490 495
 Tyr Pro Asn Ser Thr Leu Ser Val Gly Leu Ser Leu Asp Gly Pro Val
 500 505 510
 Thr Glu Ala Asp Ala Arg Val Gly Glu Leu Leu Asp Ala Leu Ala Val
 515 520 525
 Tyr Pro Arg Pro Val Phe Leu Arg Ile Gly Pro Glu Phe Asp Leu Ala
 530 535 540
 Ala Ser Gly Gln Gly Pro Glu Glu Tyr Val Ala Ala Trp Lys Thr Leu
 545 550 555 560
 His Asn Glu Ile Gln Ala Arg Gly Ser Ser Asn Ile Ala Leu Val Trp
 565 570 575
 His Ser Ala Ala Ala Cys Glu Ser Pro Phe Gly Gly His Pro Leu Glu
 580 585 590
 Ala Trp Tyr Pro Gly Asp Glu Phe Val Asp Trp Val Ala Val Ser Arg
 595 600 605
 Thr Ala Gln Ser Ala Asp Cys Glu Gly Gln Ser Val Glu Ala Val Leu
 610 615 620
 Gln Phe Ala Arg Glu Arg Tyr Lys Pro Val Val Leu Val Ala Ser Pro
 625 630 635 640
 Ala Glu Asp Ile Phe Glu Phe Val Tyr Ala Asn Asn Asp Val Ile Arg
 645 650 655
 Ala Leu Leu Tyr Leu Asn Thr Glu Pro Gly Leu Phe Asp Thr Pro Glu
 660 665 670
 Phe Leu Ser Gly Trp Lys Ala Glu Ile Gly Gln Gln Phe Trp Leu Arg
 675 680 685
 Gly Gly Pro Ala Leu Phe Ser Thr Leu Gly Leu Asp Glu
 690 695 700

<210> 7
 <211> 1539
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 7
 atggcacgtt taatcaccta ttgcttgatc ggcgtcttac tcgtgatgcc agtccttgcc 60
 gcttgacgca cagcacctac gccaacgctg atgagccagc caacttccac gccgcaaccg 120
 gccctgcaac cgacgccacc accgacgagc gtcccccggt cgatcggggc gtttgagtcc 180
 ggtcagtatc gtaatctctt cacggaatta ctgggcaaga gcgaggccga gattcagcag 240

aagatcgatc	aggcgtgggc	gcagttgttc	tacggcgaca	acgacacgca	gcgcgtttac	300
tatcccgtgg	gtcgcgacag	ggcctacatc	aaagacatcg	gcaacaatga	tgtgcgcagt	360
gaggggtatgt	cgtacggtat	gatgctggcg	gtgcagctgg	acaagcagga	agagttcaac	420
aaattgtgga	agtgggcgca	cacctatatg	ctgcaaaagg	atggcccgtg	caaaggctat	480
tttgctgggc	atgccaatga	gaacggtgaa	cagctggatg	cggttcccgc	ctccgatggc	540
gaagagtggg	ttgtcatggc	actgctcttc	gcggcaaatc	gctggggcaa	cggtgaaggc	600
atctttaatt	atcaggccga	ggcgcagaag	atcctggatg	tgatgctgca	taagagcgaa	660
gaggacaacg	gtctcgccac	cagcatgttc	gatccggaca	cgaagcaggt	ggtgtttgtg	720
ccggccgggc	gccaggccac	attcaccgat	ccgtcttatc	acttgcccgc	gttctatgaa	780
ctgtgggcgc	gtgggctga	caaggataac	gatttttggg	aagaagcggc	gcaggccagc	840
cgcgaaatttt	ggaagaaggc	ggcgcatccg	gaaacggggc	tgatgtctga	ctacgccgag	900
tttgacggca	gaccccaggc	cgattctgaa	cacaaggatt	ttcgctatga	cgcgttccgt	960
gtggcgctcca	atgtggcgct	cgattgggccc	tggttcgccc	ccgatccgtg	ggaggtggaa	1020
cacagcaaat	ggttgttggg	tttcttccgt	tcacaaggca	tggataagta	tccgagtcta	1080
tacaacatcg	atggcacgcc	gttatccact	aatcgctcgc	cggttttgat	cgccatgaac	1140
gccacagctg	gactcgcgcc	tgatccggaa	aagagcaagg	actttgtgca	ggcgctatgg	1200
gatctggaaa	ttcccagcgg	acaatggcgc	tattacgatg	gggtgctgta	tttcttgccg	1260
ctgttgcaag	ccagcggcaa	ctatcgcatc	tacacgcccg	atatgcccaa	ggtggtgcgg	1320
cccacacctg	cgcccgatcc	gatcacgcaa	gcgaaatttg	cacccggcga	tgacgcggtg	1380
ctgttccagt	tggaaacaga	tgactctgac	gaatatgtga	cggcgacggg	ctttgaagccg	1440
ggcggcgatg	tgttgaacac	tactttggac	agcgcctctt	ttgacgcacc	actgcctgac	1500
agcgctctgc	tgatcggatt	ggacgtcagc	gatcaataa			1539

<210> 8
 <211> 512
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(57)

<400> 8

Met	Ala	Arg	Leu	Ile	Thr	Tyr	Cys	Leu	Ile	Gly	Val	Leu	Leu	Val	Met
1				5					10					15	
Pro	Val	Leu	Ala	Ala	Cys	Ser	Thr	Ala	Pro	Thr	Pro	Thr	Leu	Met	Ser
			20					25					30		
Gln	Pro	Thr	Ser	Thr	Pro	Gln	Pro	Ala	Leu	Gln	Pro	Thr	Pro	Pro	Pro
			35				40					45			
Thr	Ser	Val	Pro	Arg	Ser	Ile	Gly	Ala	Phe	Glu	Ser	Gly	Gln	Tyr	Arg
			50			55					60				
Asn	Leu	Phe	Thr	Glu	Leu	Leu	Gly	Lys	Ser	Glu	Ala	Glu	Ile	Gln	Gln
					70					75				80	
Lys	Ile	Asp	Gln	Ala	Trp	Ala	Gln	Leu	Phe	Tyr	Gly	Asp	Asn	Asp	Thr
			85					90					95		
Gln	Arg	Val	Tyr	Tyr	Pro	Val	Gly	Arg	Asp	Arg	Ala	Tyr	Ile	Lys	Asp
			100				105						110		
Ile	Gly	Asn	Asn	Asp	Val	Arg	Ser	Glu	Gly	Met	Ser	Tyr	Gly	Met	Met
		115				120						125			
Leu	Ala	Val	Gln	Leu	Asp	Lys	Gln	Glu	Glu	Phe	Asn	Lys	Leu	Trp	Lys
		130				135					140				
Trp	Ala	His	Thr	Tyr	Met	Leu	Gln	Lys	Asp	Gly	Pro	Tyr	Lys	Gly	Tyr
					150					155					160
Phe	Ala	Trp	His	Ala	Asn	Glu	Asn	Gly	Glu	Gln	Leu	Asp	Ala	Gly	Pro
			165					170						175	
Ala	Ser	Asp	Gly	Glu	Glu	Trp	Phe	Val	Met	Ala	Leu	Leu	Phe	Ala	Ala
			180				185						190		
Asn	Arg	Trp	Gly	Asn	Gly	Glu	Gly	Ile	Phe	Asn	Tyr	Gln	Ala	Glu	Ala
		195				200						205			
Gln	Lys	Ile	Leu	Asp	Val	Met	Leu	His	Lys	Ser	Glu	Glu	Asp	Asn	Gly
		210				215					220				
Leu	Ala	Thr	Ser	Met	Phe	Asp	Pro	Asp	Thr	Lys	Gln	Val	Val	Phe	Val
					230					235					240
Pro	Ala	Gly	Arg	Gln	Ala	Thr	Phe	Thr	Asp	Pro	Ser	Tyr	His	Leu	Pro
			245					250						255	
Ala	Phe	Tyr	Glu	Leu	Trp	Ala	Arg	Trp	Ala	Asp	Lys	Asp	Asn	Asp	Phe
			260					265					270		

Trp Lys Glu Ala Ala Gln Ala Ser Arg Glu Phe Trp Lys Lys Ala Ala
 275 280 285
 His Pro Glu Thr Gly Leu Met Ser Asp Tyr Ala Glu Phe Asp Gly Arg
 290 295 300
 Pro Gln Ala Asp Ser Glu His Lys Asp Phe Arg Tyr Asp Ala Phe Arg
 305 310 315 320
 Val Ala Ser Asn Val Ala Leu Asp Trp Ala Trp Phe Ala Ala Asp Pro
 325 330 335
 Trp Glu Val Glu Gln Ser Asn Arg Leu Leu Asp Phe Phe Arg Ser Gln
 340 345 350
 Gly Met Asp Lys Tyr Pro Ser Leu Tyr Asn Ile Asp Gly Thr Pro Leu
 355 360 365
 Ser Thr Asn Arg Ser Pro Gly Leu Ile Ala Met Asn Ala Thr Ala Gly
 370 375 380
 Leu Ala Ala Asp Pro Glu Lys Ser Lys Asp Phe Val Gln Ala Leu Trp
 385 390 395 400
 Asp Leu Glu Ile Pro Ser Gly Gln Trp Arg Tyr Tyr Asp Gly Val Leu
 405 410 415
 Tyr Phe Leu Ala Leu Leu Gln Ala Ser Gly Asn Tyr Arg Ile Tyr Thr
 420 425 430
 Pro Asp Met Pro Lys Val Val Arg Pro Thr Pro Thr Pro Asp Pro Ile
 435 440 445
 Thr Gln Ala Lys Phe Ala Pro Gly Asp Asp Ala Val Leu Phe Ser Val
 450 455 460
 Glu Thr Asp Ala Leu Asp Glu Tyr Val Thr Ala Thr Gly Phe Glu Pro
 465 470 475 480
 Gly Gly Val Met Leu Asn Thr Thr Leu Asp Ser Ala Ser Phe Asp Ala
 485 490 495
 Pro Leu Pro Asp Ser Ala Leu Leu Ile Gly Leu Asp Val Ser Asp Gln
 500 505 510

<210> 9
 <211> 1311
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 9
 atgtttccac gtctttcacc aagccgcttc aggcaagtta ccttaacctt gctcacgctc 60
 ggccttggtg cactgaccgg ttgtgcaggt aacagcaagc cggatgcaga caccagtact 120
 gctggtgccg ttgctaccgg cgagtaccgc aatctgtttg ccgaaatcgg aaaaagcgaa 180
 atagacatcc agcgcaaaat tgacgagggc tttcagcact tgttttatgg cgacgcgaaa 240
 gatgcagctg tctactatca agcgggtgga aacgagaatg gtccactcgc atatgtttac 300
 gatgtgaaca gcaatgacgt gcgctcagaa ggcattgagct acggcatgat gattactgtt 360
 caaatggaca aaaaagccga gttcgatgca atctggaact gggcgaaaac ctatatgtat 420
 caagactccc ccacgcatcc agcgtttggt tactttgcct ggtccatgcg ccgcgatggg 480
 gtcgccaatg acgatatgcc agcggcagat ggcgaggaat atttcgtgac cgctctctat 540
 ttcgccgccg cccgctgggg taatggcgaa ggtattttca actaccaaca ggaagcggac 600
 accattttga gccgcatgcg ccaccgccag gtgatcaccg gcccaccaa tcgcggagta 660
 atgactgcga ccaatctggt ccaccgggaa gaggcgcaag tgcgcttcac gcccgacatc 720
 aataatgctg atcatacaga cgcgctttac catctgccct cgttctatga aatttgggca 780
 cgtgtcgcgc cgcaagaaga tcgcgcgttt tgggccaag cgccgatgt gagccgcgac 840
 tattttgcc aagccgccca ccctgtcact gcgttaacac cggactacgg taattttgat 900
 ggcaccccg gggcggcac ctggcgccg gagtcggtag attttcgata cgtgcctgg 960
 cgttccgtca tgaactgggt catggactat ccctgggtgg gcaaagattc aggcgcacct 1020
 gcgcgcagtg ataaattact cgcgttcttc gaaaccagg aaggcaaat gaaccacctc 1080
 tatagcctgg atggcaaacc gctgggtggt ggaccgaccc tcggcctaatt ttccatgaat 1140
 gcaacggcag ctatggcagc tactgatccc cgctggcaca attttgtgga aaagctctgg 1200
 caacaacaac cccccacagg gcaataccgg tactacgacg gtgttctata cctgatggcg 1260
 ctgctacatt gcgctgggga gtacaaagcg tggatccccg acggggaata a 1311

<210> 10
 <211> 436
 <212> PRT
 <213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(36)

<400> 10

```

Met Phe Pro Arg Leu Ser Pro Ser Arg Phe Arg Gln Val Thr Leu Thr
 1      5      10      15
Leu Leu Thr Leu Gly Leu Val Ser Leu Thr Gly Cys Ala Gly Asn Ser
 20      25      30
Lys Pro Asp Ala Asp Thr Ser Thr Ala Gly Ala Val Ala Thr Gly Glu
 35      40      45
Tyr Arg Asn Leu Phe Ala Glu Ile Gly Lys Ser Glu Ile Asp Ile Gln
 50      55      60
Arg Lys Ile Asp Glu Ala Phe Gln His Leu Phe Tyr Gly Asp Ala Lys
 65      70      75      80
Asp Ala Ala Val Tyr Tyr Gln Ala Gly Gly Asn Glu Asn Gly Pro Leu
 85      90      95
Ala Tyr Val Tyr Asp Val Asn Ser Asn Asp Val Arg Ser Glu Gly Met
 100      105      110
Ser Tyr Gly Met Met Ile Thr Val Gln Met Asp Lys Lys Ala Glu Phe
 115      120      125
Asp Ala Ile Trp Asn Trp Ala Lys Thr Tyr Met Tyr Gln Asp Ser Pro
 130      135      140
Thr His Pro Ala Phe Gly Tyr Phe Ala Trp Ser Met Arg Arg Asp Gly
 145      150      155      160
Val Ala Asn Asp Asp Met Pro Ala Pro Asp Gly Glu Glu Tyr Phe Val
 165      170      175
Thr Ala Leu Tyr Phe Ala Ala Ala Arg Trp Gly Asn Gly Glu Gly Ile
 180      185      190
Phe Asn Tyr Gln Gln Glu Ala Asp Thr Ile Leu Ser Arg Met Arg His
 195      200      205
Arg Gln Val Ile Thr Gly Pro Thr Asn Arg Gly Val Met Thr Ala Thr
 210      215      220
Asn Leu Phe His Pro Glu Glu Ala Gln Val Arg Phe Thr Pro Asp Ile
 225      230      235      240
Asn Asn Ala Asp His Thr Asp Ala Ser Tyr His Leu Pro Ser Phe Tyr
 245      250      255
Glu Ile Trp Ala Arg Val Ala Pro Gln Glu Asp Arg Ala Phe Trp Ala
 260      265      270
Lys Ala Ala Asp Val Ser Arg Asp Tyr Phe Ala Lys Ala Ala His Pro
 275      280      285
Val Thr Ala Leu Thr Pro Asp Tyr Gly Asn Phe Asp Gly Thr Pro Trp
 290      295      300
Ala Ala Ser Trp Arg Pro Glu Ser Val Asp Phe Arg Tyr Asp Ala Trp
 305      310      315      320
Arg Ser Val Met Asn Trp Ser Met Asp Tyr Ala Trp Trp Gly Lys Asp
 325      330      335
Ser Gly Ala Pro Ala Arg Ser Asp Lys Leu Leu Ala Phe Phe Glu Thr
 340      345      350
Gln Glu Gly Lys Met Asn His Leu Tyr Ser Leu Asp Gly Lys Pro Leu
 355      360      365
Gly Gly Gly Pro Thr Leu Gly Leu Ile Ser Met Asn Ala Thr Ala Ala
 370      375      380
Met Ala Ala Thr Asp Pro Arg Trp His Asn Phe Val Glu Lys Leu Trp
 385      390      395      400
Gln Gln Gln Pro Pro Thr Gly Gln Tyr Arg Tyr Tyr Asp Gly Val Leu
 405      410      415
Tyr Leu Met Ala Leu Leu His Cys Ala Gly Glu Tyr Lys Ala Trp Ile
 420      425      430
Pro Asp Gly Glu
 435

```

<210> 11

<211> 1224

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 11

atgcggaacg	tcgtgcgtaa	accattgaca	atcggactcg	ctttaacact	attattgccc	60
atgggaatga	cggcaacatc	agcgaagaat	gcagattcct	atgcgaaaaa	acctcacatc	120
agcgcatatga	atgccccaca	attggatcaa	cgctacaaaa	acgagttcac	gattgggtgcg	180
gcagtagaac	cttatcaact	acaaaatgaa	aaagacgtac	aaatgctaaa	gcgccacttc	240
aacagcattg	ttgccgagaa	cgtaatgaaa	ccgatcagca	ttcaacctga	ggaaggaaaa	300
ttcaattttg	aacaagcgga	tcgaattgtg	aagttcgcta	aggcaaatgg	catggatatt	360
cgcttccata	cactcgtttg	gcacagccaa	gtacctcaac	ggttctttct	tgacaaggaa	420
ggtaagccaa	tggtcaatga	aacagatcca	gtgaaacgtg	aacaaaataa	acaactgctg	480
ttaaaacgac	ttgaaactca	tattaaaacg	atcgtcgagc	ggtacaaaga	tgacattaag	540
tactgggacg	ttgtaaatga	ggttgtgggg	gacgacggaa	aactgcgcaa	ctctccatgg	600
tatcaaatcg	ccggcatcga	ttatatataa	gtggcattcc	aagcagctag	aaaatatggc	660
ggagacaaca	ttaagcttta	catgaatgat	tacaatacac	aagtcgaacc	gaagcgaacc	720
gctctttaca	atttagtcaa	acaactgaaa	gaagagggtg	ttccgatcga	cggcatcggc	780
catcaatccc	acatccaaat	cggctggcct	tctgaagcag	aaatcgagaa	aacgattaac	840
atgttcgccg	ctttcggttt	agacaaccaa	atcactgagc	ttgatgtgag	catgtacggt	900
tgccgcgcgc	gcgcttacc	gacgtatgac	gccattccaa	aacaaaagtt	tttggatcag	960
gcagcgcgct	atgatcgttt	gttcaaaactg	tatgagaagt	tgagcgataa	aattagcaac	1020
gtcaccttct	ggggcatcgc	cgacaatcat	acgtggctcg	acagccgtgc	ggatgtgtac	1080
tatgacgcca	acgggaatgt	tgtggttgac	ccgaacgctc	cgtacgcaa	agtggaaaaa	1140
gggaaaggaa	aagatgcgcc	gttcgttttt	ggaccggatt	acaaagtcaa	acccgcatat	1200
tgggctatta	ttgaccacaa	atag				1224

<210> 12

<211> 407

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(28)

<400> 12

Met	Arg	Asn	Val	Val	Arg	Lys	Pro	Leu	Thr	Ile	Gly	Leu	Ala	Leu	Thr
1				5					10					15	
Leu	Leu	Leu	Pro	Met	Gly	Met	Thr	Ala	Thr	Ser	Ala	Lys	Asn	Ala	Asp
			20					25					30		
Ser	Tyr	Ala	Lys	Lys	Pro	His	Ile	Ser	Ala	Leu	Asn	Ala	Pro	Gln	Leu
		35					40					45			
Asp	Gln	Arg	Tyr	Lys	Asn	Glu	Phe	Thr	Ile	Gly	Ala	Ala	Val	Glu	Pro
	50					55					60				
Tyr	Gln	Leu	Gln	Asn	Glu	Lys	Asp	Val	Gln	Met	Lys	Arg	His	Phe	
65					70					75				80	
Asn	Ser	Ile	Val	Ala	Glu	Asn	Val	Met	Lys	Pro	Ile	Ser	Ile	Gln	Pro
			85					90						95	
Glu	Glu	Gly	Lys	Phe	Asn	Phe	Glu	Gln	Ala	Asp	Arg	Ile	Val	Lys	Phe
			100					105					110		
Ala	Lys	Ala	Asn	Gly	Met	Asp	Ile	Arg	Phe	His	Thr	Leu	Val	Trp	His
		115					120					125			
Ser	Gln	Val	Pro	Gln	Arg	Phe	Phe	Leu	Asp	Lys	Glu	Gly	Lys	Pro	Met
		130				135					140				
Val	Asn	Glu	Thr	Asp	Pro	Val	Lys	Arg	Glu	Gln	Asn	Lys	Gln	Leu	Leu
145					150					155				160	
Leu	Lys	Arg	Leu	Glu	Thr	His	Ile	Lys	Thr	Ile	Val	Glu	Arg	Tyr	Lys
			165						170					175	
Asp	Asp	Ile	Lys	Tyr	Trp	Asp	Val	Val	Asn	Glu	Val	Val	Gly	Asp	Asp
			180				185						190		
Gly	Lys	Leu	Arg	Asn	Ser	Pro	Trp	Tyr	Gln	Ile	Ala	Gly	Ile	Asp	Tyr
		195				200						205			
Ile	Lys	Val	Ala	Phe	Gln	Ala	Ala	Arg	Lys	Tyr	Gly	Gly	Asp	Asn	Ile
	210					215						220			
Lys	Leu	Tyr	Met	Asn	Asp	Tyr	Asn	Thr	Glu	Val	Glu	Pro	Lys	Arg	Thr
225					230					235				240	
Ala	Leu	Tyr	Asn	Leu	Val	Lys	Gln	Leu	Lys	Glu	Glu	Gly	Val	Pro	Ile
			245						250					255	

Asp Gly Ile Gly His Gln Ser His Ile Gln Ile Gly Trp Pro Ser Glu
 260 265 270
 Ala Glu Ile Glu Lys Thr Ile Asn Met Phe Ala Ala Phe Gly Leu Asp
 275 280 285
 Asn Gln Ile Thr Glu Leu Asp Val Ser Met Tyr Gly Trp Pro Pro Arg
 290 295 300
 Ala Tyr Pro Thr Tyr Asp Ala Ile Pro Lys Gln Lys Phe Leu Asp Gln
 305 310 315 320
 Ala Ala Arg Tyr Asp Arg Leu Phe Lys Leu Tyr Glu Lys Leu Ser Asp
 325 330 335
 Lys Ile Ser Asn Val Thr Phe Trp Gly Ile Ala Asp Asn His Thr Trp
 340 345 350
 Leu Asp Ser Arg Ala Asp Val Tyr Tyr Asp Ala Asn Gly Asn Val Val
 355 360 365
 Val Asp Pro Asn Ala Pro Tyr Ala Lys Val Glu Lys Gly Lys Lys
 370 375 380
 Asp Ala Pro Phe Val Phe Gly Pro Asp Tyr Lys Val Lys Pro Ala Tyr
 385 390 395 400
 Trp Ala Ile Ile Asp His Lys
 405

<210> 13
 <211> 1053
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 13
 atgaaagacg cgctccagtg ctctcccctt ttcaaagcct atgaaaaata cttccgcatac 60
 ggcgcggcgg ttagcagctt catgaccttt gatcccgctt accgcgccct gatccgccgc 120
 cattacaatt cctgacggc ggacaaccag atgaagccgg aaagcgtgtt ggatcgcacc 180
 gcgaccctgg cgaagggcga cctgctccac gctgcggtgg atttcaccg tgtggacgcg 240
 ctgatgtact ttgcacggga caacgggatc cccatgcggt atcacaccct ggctggcac 300
 aaccagacgc cccgctgggt cttcggaag gactggagcg acgcggaaag cgccgaaccc 360
 gcctcaaagg aaacctgct tgcccgtctg gaaaactata tcctggatgt catgaacct 420
 gtgaatacca agtttcccgg tctggtttac acctgggacg tggtaaacga agccattgag 480
 ccagagctga aagccccggg attgtaccgg acctggagcc cctggttcaa aacctgcgga 540
 gaagatttcc tctttaccgc tttccgggccc gccgcgaagg gacaggcgcc cggtcagacc 600
 ctttgcata acgactataa cgccttcgag cccgtcaagc gggacgcgat tatcgatctg 660
 ctgaagaagc tgcaggcgga aaacctggtg gataccatgg gtatgcagg gcattatgtc 720
 atggactgga tgaacatctc gctctgcgaa gaggccgcc gcgcctatgc cgccctgggc 780
 ctgaagggtcc aggtcaccga gctggatatc cactgcaaca gcgacgatga agccacagc 840
 caaaagctgg cgcagcttta cggcgattat ttccgcatgc tgaagaagct gaaggaggaa 900
 ggcgtcgaca tcgaagccgt caccttctgg ggctcaccg accaggacag ctggctcacc 960
 ggtttcgta aagagacaag ctatcccctc ctctcgacc gcgccaagca ggccaaggat 1020
 gcctatgacg ccgtcatgaa agccgcggaa taa 1053

<210> 14
 <211> 350
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 14
 Met Lys Asp Ala Leu Gln Cys Ser Pro Leu Phe Lys Ala Tyr Glu Lys
 1 5 10 15
 Tyr Phe Arg Ile Gly Ala Ala Val Ser Phe Met Thr Phe Asp Pro
 20 25 30
 Ala Tyr Arg Ala Leu Ile Arg Arg His Tyr Asn Ser Leu Thr Ala Asp
 35 40 45
 Asn Gln Met Lys Pro Glu Ser Val Leu Asp Arg Thr Ala Thr Leu Ala
 50 55 60
 Lys Gly Asp Leu Leu His Ala Ala Val Asp Phe Thr Arg Val Asp Ala
 65 70 75 80
 Leu Met Tyr Phe Ala Arg Asp Asn Gly Ile Pro Met Arg Tyr His Thr
 Page 11

Leu	Ala	Trp	His	85	Asn	Gln	Thr	Pro	Arg	90	Trp	Phe	Phe	Ala	Lys	95	Asp	Trp
Ser	Asp	Ala	100	Glu	Ser	Ala	Glu	Pro	105	Ala	Ser	Lys	Glu	Thr	110	Met	Leu	Ala
Arg	Leu	115	Glu	Asn	Tyr	Ile	Leu	120	Asp	Val	Met	Asn	His	125	Val	Asn	Thr	Lys
130	Phe	Pro	Gly	Leu	Val	Tyr	135	Thr	Trp	Asp	Val	Val	140	Asn	Glu	Ala	Ile	Glu
145	Pro	Glu	Leu	Lys	Ala	150	Pro	Gly	Leu	Tyr	Arg	155	Thr	Trp	Ser	Pro	Trp	Phe
Lys	Thr	Cys	Gly	165	Glu	Asp	Phe	Leu	Phe	170	Thr	Ala	Phe	Arg	Ala	Ala	Arg	
Lys	Gly	Gln	Ala	180	Pro	Gly	Gln	Thr	Leu	185	Cys	Tyr	Asn	205	Asp	Tyr	Asn	Ala
195	Phe	Glu	Pro	Val	Lys	Arg	Asp	Ala	Ile	200	Ile	Asp	Leu	220	Leu	Lys	Lys	Leu
210	Gln	Ala	Glu	Asn	Leu	215	Val	Asp	Thr	Met	Gly	Met	Gln	235	Gly	His	Tyr	Val
225	Met	Asp	Trp	Met	Asn	230	Ile	Ser	Leu	Cys	Glu	Glu	Ala	255	Ala	Arg	Ala	Tyr
245	Ala	Ala	Leu	Gly	Leu	Lys	Val	Gln	Val	265	Thr	Glu	Leu	270	Asp	Ile	His	Cys
260	Asn	Ser	Asp	Asp	Glu	Ala	His	Ser	Gln	280	Lys	Leu	Ala	285	Gln	Leu	Tyr	Gly
275	Asp	Tyr	Phe	Ala	Met	Leu	Lys	Lys	Leu	295	Lys	Glu	Glu	300	Gly	Val	Asp	Ile
290	Glu	Ala	Val	Thr	Phe	Trp	Gly	Val	Thr	310	Asp	Gln	Asp	315	Ser	Trp	Leu	Thr
305	Gly	Phe	Arg	Lys	Glu	Thr	Ser	Tyr	Pro	325	Leu	Leu	Phe	330	Asp	Arg	Ala	Lys
325	Gln	Ala	Lys	Asp	Ala	Tyr	Asp	Ala	Val	335	Met	Lys	Ala	345	Ala	Glu		
340																		

<210> 15
 <211> 1110
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 15	atgaaacgctc	ctctagtcaa	tctcctgaca	accgcctgcc	tcctcgttgc	cgcaaagtct	60
	gcagaaccca	ccctccgcga	agcctacgaa	aagcactttg	ccgtgggtgt	cgactcaat	120
	accgctcaag	tgactggtcg	aaacaaagcc	gcaggcgaac	tcgccgcgaa	gcagttcaat	180
	tccatcaccg	ctgagaatga	catgaagtgg	caatcgcttc	atccagagct	cgatacctac	240
	cgctttgaat	cggccgatgc	ctatatcgac	tttgccaaaa	agaatgagat	ggaagtcata	300
	ggccacactc	tcgtctggca	cagccagacc	cctcagtggg	tgttccaagg	cgacgatggc	360
	aaaccgcgca	cacgggaaga	acttctcaag	cggatgcgcg	atcacattca	caaggctgcc	420
	ggccgataca	agggtaaggt	caagggctgg	gacgtcgtca	atgaggcgct	ctccgacgga	480
	ggtcaggaca	ttctacgcga	atctccgtgg	cggcgaatca	tcggagacga	tttcatcgat	540
	cacgctttcc	gctacgcccg	cgaagccgac	ccaaaggcag	aactttacta	caacgactac	600
	aacctcgaaa	tccttcgcaa	acgcgagaac	tgcatcaagc	tcgtcaaggg	catgcttgag	660
	cgcggcgtcc	ccatcgacgg	cattggaacg	caatcccaat	ttcagcttgg	cttcccatcg	720
	ctggaagatg	tcgagaccac	gattgaagag	tttggaatac	tcggccttaa	ggtcatgatt	780
	accgaactcg	atgtggatgt	cctccctcgc	aataaccagc	gcgtcgccga	catcagtcag	840
	cgcgagcaag	gtagcaatcc	ctacactgag	ggcctgcccg	aggatgttca	aaagcagctt	900
	acgaacgctc	acgaagacat	cttcaagatc	tacctaaagc	accagaaaac	ggtcaccgct	960
	gtgaccttct	ggggcctcga	tgatgggtcaa	tcatgggtga	atggccttcc	tgtagagggc	1020
	cgaccaatc	accgcctact	tttcgatcgt	gaactcaaac	cgaagcccgt	tcttccagtc	1080
	ttgatagagc	tcggcaagaa	gaagcgataa				1110

<210> 16
 <211> 369
 <212> PRT
 <213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(20)

<400> 16

```

Met Lys Arg Pro Leu Val Asn Leu Leu Thr Thr Ala Cys Leu Leu Val
 1      5      10      15
Ala Ala Asn Ala Ala Glu Pro Thr Leu Arg Glu Ala Tyr Glu Lys His
 20      25      30
Phe Ala Val Gly Val Ala Leu Asn Thr Ala Gln Val Thr Gly Arg Asn
 35      40      45
Lys Ala Ala Gly Glu Leu Ala Ala Lys Gln Phe Asn Ser Ile Thr Ala
 50      55      60
Glu Asn Asp Met Lys Trp Gln Ser Leu His Pro Glu Leu Asp Thr Tyr
 65      70      75      80
Arg Phe Glu Ser Ala Asp Ala Tyr Ile Asp Phe Ala Lys Lys Asn Glu
 85      90      95
Met Glu Val Ile Gly His Thr Leu Val Trp His Ser Gln Thr Pro Gln
 100     105     110
Trp Val Phe Gln Gly Asp Asp Gly Lys Pro Ala Thr Arg Glu Glu Leu
 115     120     125
Leu Lys Arg Met Arg Asp His Ile His Lys Val Ala Gly Arg Tyr Lys
 130     135     140
Gly Lys Val Lys Gly Trp Asp Val Val Asn Glu Ala Leu Ser Asp Gly
 145     150     155     160
Gly Gln Asp Ile Leu Arg Glu Ser Pro Trp Arg Arg Ile Ile Gly Asp
 165     170
Asp Phe Ile Asp His Ala Phe Arg Tyr Ala Arg Glu Ala Asp Pro Lys
 180     185     190
Ala Glu Leu Tyr Tyr Asn Asp Tyr Asn Leu Glu Ile Pro Arg Lys Arg
 195     200     205
Glu Asn Cys Ile Lys Leu Val Lys Gly Met Leu Glu Arg Gly Val Pro
 210     215     220
Ile Asp Gly Ile Gly Thr Gln Ser His Phe Gln Leu Gly Phe Pro Ser
 225     230     235     240
Leu Glu Asp Val Glu Thr Thr Ile Glu Glu Phe Gly Lys Leu Gly Leu
 245     250     255
Lys Val Met Ile Thr Glu Leu Asp Val Asp Val Leu Pro Arg Asn Asn
 260     265     270
Pro Gly Val Ala Asp Ile Ser Gln Arg Glu Gln Gly Ser Asn Pro Tyr
 275     280     285
Thr Glu Gly Leu Pro Glu Asp Val Gln Lys Gln Leu Thr Lys Arg Tyr
 290     295     300
Glu Asp Ile Phe Lys Ile Tyr Leu Lys His Gln Lys Thr Val Thr Arg
 305     310     315     320
Val Thr Phe Trp Gly Leu Asp Asp Gly Gln Ser Trp Leu Asn Gly Phe
 325     330     335
Pro Val Arg Gly Arg Thr Asn His Pro Leu Leu Phe Asp Arg Glu Leu
 340     345     350
Lys Pro Lys Pro Val Leu Pro Val Leu Ile Glu Leu Gly Lys Lys Lys
 355     360     365
Arg

```

<210> 17

<211> 1035

<212> DNA

<213> Bacteria

<400> 17

```

atgtcccggc acgtcatcgc cctgtccgcc gccgtctgcc tcgcgggccgg cctcgccgcc      60
gcgcccgcga gcgccgagcc gcgtccccgg acgctcggcg aactggccaa gaagcaccac      120
aagtacttcg gctcggccac cgacaacccc gagttcaccg acgccgccta tctgaagctc      180
ctcggcagcg agttcgggca gaccaccccc ggcaacgcc a tgaagtggta cgccaccgaa      240
cccgcgcccg gcgtcttcga cttcaccgcg ggcgacgagg tcgtggcctt cgccaaggcc      300
catcaccaga aggtccgcgg ccacaccctc gtctggcaca gccagctccc cgctggctc      360
accgagcgca gctggaccgc cgcggaactg cgcccgtcc tcaagaatca catccagaag      420
gtggcccggc actacaaggg caaggtcatc cactgggacg tcgtcaacga ggccttcaac      480

```


gaggacggca	cctaccgcga	gtcgggtcttc	tacaagacgc	tcggccccgg	ctacatcgcc	540
gacgccctgc	gctggggccca	cgaggccgac	ccgcacgccca	agctgtacct	caacgactac	600
aacgtcgacg	ggatcggccc	caagagcgac	gcctactacc	gcctgatcaa	gcagctgaag	660
gccgacggcg	tcccgggtgga	gggcttcggc	atccaggggc	acctggcgct	ccagtacggc	720
ttccccgcgg	acgtcaagca	gaacatgcag	cgcttcgccg	acctcggcgt	cgaggtcgcg	780
gtcaccgagc	tcgacatccg	gatgaacctc	ccggcgaccc	cttcgatgct	cgccaccag	840
gccacctggg	acgccgacta	cgtaaggccc	tgcctggagg	tcaggaagtg	cgtcggcgctc	900
accatctggg	actacaccga	caagtactcg	tggatcccct	ccgtcttccc	cggtgagggc	960
gccgcgctgc	cctacgacga	gaacctggcg	ccaagccccg	cctaccacgc	gatcaggaag	1020
gtgctgggcg	gatga					1035

<210> 18
 <211> 344
 <212> PRT
 <213> Bacteria

<220>
 <221> SIGNAL
 <222> (1)...(31)

<400> 18
 Met Ser Arg His Val Ile Ala Leu Ser Ala Ala Val Cys Leu Ala Ala
 1 5 10 15
 Gly Leu Ala Ala Pro Ala Ser Ala Glu Pro Arg Pro Arg Thr Leu
 20 25 30
 Gly Glu Leu Ala Lys Lys His His Lys Tyr Phe Gly Ser Ala Thr Asp
 35 40 45
 Asn Pro Glu Phe Thr Asp Ala Ala Tyr Leu Lys Leu Leu Gly Ser Glu
 50 55 60
 Phe Gly Gln Thr Thr Pro Gly Asn Ala Met Lys Trp Tyr Ala Thr Glu
 65 70 75 80
 Pro Ala Pro Gly Val Phe Asp Phe Thr Ala Gly Asp Glu Val Val Ala
 85 90 95
 Phe Ala Lys Ala His His Gln Lys Val Arg Gly His Thr Leu Val Trp
 100 105 110
 His Ser Gln Leu Pro Ala Trp Leu Thr Glu Arg Ser Trp Thr Ala Ala
 115 120 125
 Glu Leu Arg Pro Val Leu Lys Asn His Ile Gln Lys Val Ala Arg His
 130 135 140
 Tyr Lys Gly Lys Val Ile His Trp Asp Val Val Asn Glu Ala Phe Asn
 145 150 155 160
 Glu Asp Gly Thr Tyr Arg Glu Ser Val Phe Tyr Lys Thr Leu Gly Pro
 165 170 175
 Gly Tyr Ile Ala Asp Ala Leu Arg Trp Ala His Glu Ala Asp Pro His
 180 185 190
 Ala Lys Leu Tyr Leu Asn Asp Tyr Asn Val Asp Gly Ile Gly Pro Lys
 195 200 205
 Ser Asp Ala Tyr Tyr Arg Leu Ile Lys Gln Leu Lys Ala Asp Gly Val
 210 215 220
 Pro Val Glu Gly Phe Gly Ile Gln Gly His Leu Ala Leu Gln Tyr Gly
 225 230 235 240
 Phe Pro Ala Asp Val Lys Gln Asn Met Gln Arg Phe Ala Asp Leu Gly
 245 250 255
 Val Glu Val Ala Val Thr Glu Leu Asp Ile Arg Met Asn Leu Pro Ala
 260 265 270
 Thr Pro Ser Met Leu Ala Thr Gln Ala Thr Trp Tyr Ala Asp Tyr Val
 275 280 285
 Lys Ala Cys Leu Glu Val Arg Lys Cys Val Gly Val Thr Ile Trp Asp
 290 295 300
 Tyr Thr Asp Lys Tyr Ser Trp Ile Pro Ser Val Phe Pro Gly Glu Gly
 305 310 315 320
 Ala Ala Leu Pro Tyr Asp Glu Asn Leu Ala Pro Lys Pro Ala Tyr His
 325 330 335
 Ala Ile Arg Lys Val Leu Gly Gly
 340

<210> 19
 <211> 1152
 <212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 19

atgaagatgt	taaaaactat	tgttgtggct	gtagcagcct	tactatccag	tcctactgct	60
tcagccactt	tacagaacct	gaagcgggct	cctgattcat	tgaccttgaa	agatgcattt	120
gagggtaagt	tttatatagg	aacagcatta	aaccttgatc	agatatggga	gcgcgatcag	180
gctgcggctc	cggtgggtcaa	aacgcagttc	aactccatag	ttgctgagaa	ttgtatgaaa	240
agtatgtttt	tgaaccaag	ggaagggtgag	tttgatttta	gggatgcgga	ccgttttgct	300
gcgtttggag	aaaaaataa	aatgcaaatt	atcggtcata	cgctgatttg	gcattcgcag	360
acaccagctt	ggttttttgt	cgataaaaaat	gggaaagagg	tcacccgaga	ggtacttatc	420
gagcgcgatg	ggaagcatat	acaaaccggt	gtttcccgct	ataaggggaag	ggtgtttggt	480
tgggatgttg	tgaacgaagc	catattggat	aatggagaat	ggcgtaaaag	caaatttctac	540
cagattatcg	ggccacaatt	tattgaattg	gccttcaa	ttgcgcgatga	cgcatgcca	600
aatgcagaat	tatattataa	cgattattca	actgctatcc	ccgaaaaaag	aaaggggatt	660
atgcgcgatg	tgcagcaggt	aaaggctgcc	ggtgggcagg	tcactggaat	tggatgcag	720
gaacacaacg	cattggacaa	tccaccggct	gatgaagtcg	aaaaaaccat	actcggattt	780
gcaagccttg	gtgcgaagg	aatgggttac	gaaatggata	tttcggctcct	gccgcagtga	840
cgtcccaata	tgggcgcaga	aataggggag	cgtcatgcct	acagtaaagc	gatgaatccg	900
tacgaaaaag	gacttcctgt	aacgaaaatg	aacgagttgg	gagcgagata	tgtagcgttt	960
tttaatttat	atctcaaaca	tcgggataaa	atatcgcgtg	tgacattgtg	gggtgttggc	1020
gatggagatt	catggaagaa	tggttggcct	attccgggac	gtacagacta	tccattgtta	1080
ttcgatcgga	attaccaacc	caaacccttt	gtaaaagata	ttattgcgtt	gactcaaaaa	1140
aaaaagaaat	aa					1152

<210> 20

<211> 383

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(29)

<400> 20

Met	Lys	Met	Leu	Lys	Thr	Ile	Val	Val	Ala	Val	Ala	Ala	Leu	Leu	Ser	
1				5					10					15		
Ser	Pro	Thr	Ala	Ser	Ala	Thr	Leu	Gln	Asn	Leu	Lys	Arg	Ala	Pro	Asp	
			20					25					30			
Ser	Leu	Thr	Leu	Lys	Asp	Ala	Phe	Glu	Gly	Lys	Phe	Tyr	Ile	Gly	Thr	
			35				40					45				
Ala	Leu	Asn	Leu	Asp	Gln	Ile	Trp	Glu	Arg	Asp	Gln	Ala	Ala	Val	Ala	
	50				55						60					
Val	Val	Lys	Thr	Gln	Phe	Asn	Ser	Ile	Val	Ala	Glu	Asn	Cys	Met	Lys	
65				70					75					80		
Ser	Met	Phe	Leu	Gln	Pro	Arg	Glu	Gly	Glu	Phe	Asp	Phe	Arg	Asp	Ala	
			85					90						95		
Asp	Arg	Phe	Val	Ala	Phe	Gly	Glu	Lys	Asn	Lys	Met	Gln	Ile	Ile	Gly	
			100					105					110			
His	Thr	Leu	Ile	Trp	His	Ser	Gln	Thr	Pro	Ala	Trp	Phe	Phe	Val	Asp	
		115					120					125				
Lys	Asn	Gly	Lys	Glu	Val	Thr	Arg	Glu	Val	Leu	Ile	Glu	Arg	Met	Arg	
	130					135					140					
Lys	His	Ile	Gln	Thr	Val	Val	Ser	Arg	Tyr	Lys	Gly	Arg	Val	Phe	Gly	
145					150					155				160		
Trp	Asp	Val	Val	Asn	Glu	Ala	Ile	Leu	Asp	Asn	Gly	Glu	Trp	Arg	Lys	
			165					170						175		
Ser	Lys	Phe	Tyr	Gln	Ile	Ile	Gly	Pro	Gln	Phe	Ile	Glu	Leu	Ala	Phe	
			180					185					190			
Lys	Phe	Ala	His	Asp	Ala	Asp	Pro	Asn	Ala	Glu	Leu	Tyr	Tyr	Asn	Asp	
		195					200					205				
Tyr	Ser	Thr	Ala	Ile	Pro	Glu	Lys	Arg	Lys	Gly	Ile	Met	Arg	Met	Val	
	210					215					220					
Gln	Gln	Val	Lys	Ala	Ala	Gly	Gly	Gln	Val	Thr	Gly	Ile	Gly	Met	Gln	
225				230					235					240		

Glu His Asn Ala Leu Asp Asn Pro Pro Val Asp Glu Val Glu Lys Thr
 Ile Leu Gly Phe Ala Ser Leu Gly Ala Lys Val Met Val Thr Glu Met
 Asp Ile Ser Val Leu Pro His Val Arg Pro Asn Met Gly Ala Glu Ile
 Gly Glu Arg His Ala Tyr Ser Lys Ala Met Asn Pro Tyr Glu Lys Gly
 Leu Pro Val Thr Lys Met Asn Glu Leu Gly Ala Arg Tyr Val Ala Phe
 Phe Asn Leu Tyr Leu Lys His Arg Asp Lys Ile Ser Arg Val Thr Leu
 Trp Gly Val Gly Asp Gly Asp Ser Trp Lys Asn Gly Trp Pro Ile Pro
 Gly Arg Thr Asp Tyr Pro Leu Leu Phe Asp Arg Asn Tyr Gln Pro Lys
 Pro Phe Val Lys Asp Ile Ile Ala Leu Thr Gln Lys Lys Lys Lys
 370 375 380

<210> 21
 <211> 1119
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 21
 atgcggattc actggctggg gctcagctca cgcgcaagcc tgatgacggc ggcgctcctg 60
 gctgtcacag gcaccaccaa atccgaggac tcgcccga ctttgaaaga cgccttcaag 120
 gattgtttcc ggatcggggt cgcgctcaac cagcggcaat ttaccgagca agataccaac 180
 ggcgcgacgt tggatgaaac gcagttcaac gccatctcac ccgaaaacgt gatgaagtgg 240
 gcgaacattc atccccgacc cgggcccgat ggggtataact tcgagggcggc tgaccgttac 300
 gtcgagtgtt gcgagaagaa cgggaatgttc atcgtcggcc atacgctcgt ttggcacttc 360
 caaacgccgc gctgggtact ccagggcgat ggcactaacg cggcgacgcg cgagctgctg 420
 ctgcagcggg tgcgcgatca catccacacg gtcgtaggcc ggtacaaagg gcggatcaag 480
 gcttgggacg tggatcaacga agcgtgaac gaagatggca ctctgcggcg gtcgcagtgg 540
 taccggatca tcggcgaaga ctacatcgct aaggctttcg aatatgcgca tgaggccgat 600
 ccgtccgcgg aattgcgata caacgattac gccatcgaga atgagcggaa gcgcgacggc 660
 gtaatcgcg tcgtgaagaa acttcaggcg cagaaggtcc cacttggggg gctgggctcg 720
 cagacgcgat ccaacctgac ctggcctaac gccgaatcgc tggacaccgc cctcacggcc 780
 ttaccgaac tgggtatccc gatctcaatc acggaactgg atgtgaccgc ctgcgaacgc 840
 ggtcagctca accagagcgc cgaggtgtcg cagaatggac aggcggggga gggaggcgtg 900
 gtggacgggg cgaatcagaa gctcgccgag cagtacgcca acttcttccg cgtctttctg 960
 aagcatcgca aaaacattga gctcgtgacg ttttgggcg tcacggatcg tgactcctgg 1020
 cggcgcatcg gcaaaccgct gctatttaac gcagaatggc aaccaagcc ggcctttcac 1080
 gccgtcatcg ccgaggcgaa aaagatcagt gggcaatga 1119

<210> 22
 <211> 372
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 22
 Met Arg Ile His Trp Leu Gly Leu Ser Ser Arg Ala Ser Leu Met Thr
 1 5 10 15
 Ala Ala Leu Leu Ala Val Thr Gly Thr Thr Lys Ser Glu Asp Ser Pro
 20 25 30
 Ala Thr Leu Lys Asp Ala Phe Lys Asp Cys Phe Arg Ile Gly Val Ala
 35 40 45
 Leu Asn Gln Arg Gln Phe Thr Glu Gln Asp Thr Asn Gly Ala Thr Leu
 50 55 60
 Val Lys Arg Gln Phe Asn Ala Ile Ser Pro Glu Asn Val Met Lys Trp

65 Ala Asn Ile His Pro Arg Pro Gly Pro Asp Gly Tyr Asn Phe Glu Ala
 70 85 90 95
 Ala Asp Arg Tyr Val Glu Phe Gly Glu Lys Asn Gly Met Phe Ile Val
 100 105 110
 Gly His Thr Leu Val Trp His Phe Gln Thr Pro Arg Trp Val Leu Gln
 115 120 125
 Gly Asp Gly Thr Asn Ala Ala Thr Arg Glu Leu Leu Leu Gln Arg Met
 130 135 140
 Arg Asp His Ile His Thr Val Val Gly Arg Tyr Lys Gly Arg Ile Lys
 145 150 155 160
 Ala Trp Asp Val Val Asn Glu Ala Leu Asn Glu Asp Gly Thr Leu Arg
 165 170 175
 Arg Ser Gln Trp Tyr Arg Ile Ile Gly Glu Asp Tyr Ile Val Lys Ala
 180 185 190
 Phe Glu Tyr Ala His Glu Ala Asp Pro Ser Ala Glu Leu Arg Tyr Asn
 195 200 205
 Asp Tyr Ala Ile Glu Asn Glu Arg Lys Arg Asp Gly Val Ile Ala Leu
 210 215 220
 Val Lys Lys Leu Gln Ala Gln Lys Val Pro Leu Gly Gly Leu Gly Ser
 225 230 235 240
 Gln Thr His Ala Asn Leu Thr Trp Pro Asn Ala Glu Ser Leu Asp Thr
 245 250 255
 Ala Leu Thr Ala Phe Thr Glu Leu Gly Ile Pro Ile Ser Ile Thr Glu
 260 265 270
 Leu Asp Val Thr Ala Ser Gln Arg Gly Gln Leu Asn Gln Ser Ala Glu
 275 280 285
 Val Ser Gln Asn Gly Gln Ala Gly Glu Gly Gly Val Val Asp Gly Ala
 290 295 300
 Asn Gln Lys Leu Ala Glu Gln Tyr Ala Asn Phe Phe Arg Val Phe Leu
 305 310 315 320
 Lys His Arg Lys Asn Ile Glu Leu Val Thr Phe Trp Gly Val Thr Asp
 325 330 335
 Arg Asp Ser Trp Arg Arg Ile Gly Lys Pro Leu Leu Phe Asn Ala Glu
 340 345 350
 Trp Gln Pro Lys Pro Ala Phe His Ala Val Ile Ala Glu Ala Lys Lys
 355 360 365
 Ile Ser Gly Gln
 370

<210> 23
 <211> 1137
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 23
 atgaggacaa aacaagtttt taaattaacc acgctcgctt tattattaac agcagttggt 60
 agtagctggt ctgccccaaa agcggcaaaa gaagatacgc ttaaagatgc cctccaggga 120
 aaattcttta ttggtgctgc tgtaaatggt gaccaaattg caggaaagga ttctcttgca 180
 attgaagttg ttaaaaagaa ctttagctca attgtggccg agaattgcat gaaaatggaa 240
 aacatccatc ctgtaaaagg tgaatttttc ttcgatgaag ccgatgcata tgttgaattt 300
 ggcgaaaaaa acaacatgaa aatcattggt cacacattga tttggcattc acaagccgcc 360
 aaatgggcat ttgttgatga tgaaggcaaa gatgtatcgc gcgaagaatt aattgaacgg 420
 atgcgcaacc acatccatac cattgtaggc cgctataaag gtcgtgtaca tggctgggac 480
 gttgtaaatg aggctattct ggataacggc gaatggcgct agagcaaatt gtataccatt 540
 attggaccgg aatttgttca gcttgctttt gagtttgccc acgaagccga cccaacgct 600
 gaattgtatt acaacgacta caacgagtgg attccggcta aaagagacgg catttacaac 660
 atggttaagg atttaacgga caaaggcggt aaagttagtg gaattggcct acagggtcac 720
 attgctcttg actctccag catcgaactt tacgaagaag ccattgtaaa atatgcaagt 780
 ctgggtgtgc aaacaatggt taccgaactc gatatactg ttttaccatg gccatcgag 840
 caagttagac ccgatataatc ttttagtgca gagctatcaa ccgaatacaa tccatttggt 900
 aatggtttac ccgattcggt tagcgttgaa cttaccaacc gttttgccag tttcttcgag 960
 ttgtttttga aacatcagga taaaattgac cgcgttactc tatgggggtg acacgatggt 1020
 caatcatgga aaaacaactg gccatcagg ggacgtaaag attatccggt gttattcgac 1080
 aggcaatatc agtccaaacc tgcggttcag cgcataatcg aattggctaa acaataa 1137

<210> 24
 <211> 378
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(29)

<400> 24
 Met Arg Thr Lys Gln Val Phe Lys Leu Thr Thr Leu Ala Leu Leu Leu
 1 5 10 15
 Thr Ala Val Val Ser Ser Cys Ser Ala Pro Lys Ala Ala Lys Glu Asp
 20 25 30
 Thr Leu Lys Asp Ala Leu Gln Gly Lys Phe Phe Ile Gly Ala Ala Val
 35 40 45
 Asn Val Asp Gln Met Ala Gly Lys Asp Ser Leu Ala Ile Glu Val Val
 50 55 60
 Lys Lys Asn Phe Ser Ser Ile Val Ala Glu Asn Cys Met Lys Met Glu
 65 70 75 80
 Asn Ile His Pro Val Lys Gly Glu Phe Phe Phe Asp Glu Ala Asp Ala
 85 90 95
 Tyr Val Glu Phe Gly Glu Lys Asn Asn Met Lys Ile Ile Gly His Thr
 100 105 110
 Leu Ile Trp His Ser Gln Ala Ala Lys Trp Ala Phe Val Asp Asp Glu
 115 120 125
 Gly Lys Asp Val Ser Arg Glu Glu Leu Ile Glu Arg Met Arg Asn His
 130 135 140
 Ile His Thr Ile Val Gly Arg Tyr Lys Gly Arg Val His Gly Trp Asp
 145 150 155 160
 Val Val Asn Glu Ala Ile Leu Asp Asn Gly Glu Trp Arg Gln Ser Lys
 165 170 175
 Trp Tyr Thr Ile Ile Gly Pro Glu Phe Val Gln Leu Ala Phe Glu Phe
 180 185 190
 Ala His Glu Ala Asp Pro Asn Ala Glu Leu Tyr Tyr Asn Asp Tyr Asn
 195 200 205
 Glu Trp Ile Pro Ala Lys Arg Asp Gly Ile Tyr Asn Met Val Lys Asp
 210 215 220
 Leu Ile Asp Lys Gly Val Lys Val Asp Gly Ile Gly Leu Gln Gly His
 225 230 235 240
 Ile Ala Leu Asp Ser Pro Ser Ile Glu Leu Tyr Glu Glu Ala Ile Val
 245 250 255
 Lys Tyr Ala Ser Leu Gly Val Gln Thr Met Val Thr Glu Leu Asp Ile
 260 265 270
 Thr Val Leu Pro Trp Pro Ser Gln Val Thr Ala Asp Ile Ser Phe
 275 280 285
 Ser Ala Glu Leu Ser Thr Glu Tyr Asn Pro Phe Val Asn Gly Leu Pro
 290 295 300
 Asp Ser Val Ser Val Glu Leu Thr Asn Arg Phe Ala Ser Phe Phe Glu
 305 310 315 320
 Leu Phe Leu Lys His Gln Asp Lys Ile Asp Arg Val Thr Leu Trp Gly
 325 330 335
 Val His Asp Gly Gln Ser Trp Lys Asn Asn Trp Pro Ile Arg Gly Arg
 340 345 350
 Lys Asp Tyr Pro Leu Leu Phe Asp Arg Gln Tyr Gln Ser Lys Pro Ala
 355 360 365
 Val Gln Arg Ile Ile Glu Leu Ala Lys Gln
 370 375

<210> 25
 <211> 978
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 25
 gtggatccaa agaattcctt acgcgcctta gctcaaaagc gaggaattgg gtttgggacg 60
 gcagtttggg ttgagcctct gtctaacgat tgcagatata ggacggtgtt ggcgcaggag 120
 ttcaatatgg tgacgccaga gaatgagatg aagtttgagc cgacgcatcc agaacgggag 180
 cgctacgatt ttacagcagc cgataccctt gttagctttg ccaagaacca taacatgcag 240
 gtgcgcggac ataccctggt ttggcatgaa agtctccccg attggctaac gactcaaacg 300
 tggacgcgtg aggagttgat gtccatctta gaagaacaca tcaatacagt tgcgatcgc 360
 tatcgggggc aattagttgc ctgggatgtg gtgaatgaag cgatcgccaa cgataaaaac 420
 gcactcagag atacgatttg gctgcgaaca atcgggcccag agtatataga gaaggcattt 480
 cgctgggccc atgcagcccga cctcaagca cgtttatttt acaacgatta tggcggcgag 540
 gaagtggggg gaaagtctga ggccatctat ggcatgctta aagatttgct gcaacagggt 600
 gtcccgaattc acgggggttg cttgcaaattg cacgttagta taaaaaaccc tcccaatccc 660
 gaaaaagtgg cggcaaatat caagcgcctg aacgatctgg gattggaagt gcatataact 720
 gagatggatg tgaaaacctg ggatggcatc ggtacgaagc agcaacgact tgcggctcag 780
 gcacaagtgt atcgaacat gatgcagggt tgtttggaag ctgagaactg taaggcggtt 840
 tcgttgtggg gggtaagcga tcgctattct tggattcccc ggatttttaa gaagccggat 900
 gcaccactga tttttgatga tttagggcgt ccgaaacccg cttacaatgc cctgaaagaa 960
 gtcctcaagc ggcgttaa 978

<210> 26
 <211> 325
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 26
 Val Asp Pro Lys Asn Ser Leu Arg Ala Leu Ala Gln Lys Arg Gly Ile
 1 5 10 15
 Gly Phe Gly Thr Ala Val Trp Val Glu Pro Leu Ser Asn Asp Ser Arg
 20 25 30
 Tyr Arg Thr Val Leu Ala Gln Glu Phe Asn Met Val Thr Pro Glu Asn
 35 40 45
 Glu Met Lys Phe Glu Pro Thr His Pro Glu Arg Glu Arg Tyr Asp Phe
 50 55 60
 Thr Ala Ala Asp Thr Leu Val Asp Phe Ala Lys Asn His Asn Met Gln
 65 70 75 80
 Val Arg Gly His Thr Leu Val Trp His Glu Ser Leu Pro Asp Trp Leu
 85 90 95
 Thr Thr Gln Thr Trp Thr Arg Glu Glu Leu Met Ser Ile Leu Glu Glu
 100 105 110
 His Ile Asn Thr Val Val Asp Arg Tyr Arg Gly Gln Leu Val Ala Trp
 115 120 125
 Asp Val Val Asn Glu Ala Ile Ala Asn Asp Lys Asn Ala Leu Arg Asp
 130 135 140
 Thr Ile Trp Leu Arg Thr Ile Gly Pro Glu Tyr Ile Glu Lys Ala Phe
 145 150 155 160
 Arg Trp Ala His Ala Ala Asp Pro Gln Ala Arg Leu Phe Tyr Asn Asp
 165 170 175
 Tyr Gly Gly Glu Glu Val Gly Gly Lys Ser Glu Ala Ile Tyr Gly Met
 180 185 190
 Leu Lys Asp Leu Leu Gln Gln Gly Val Pro Ile His Gly Val Gly Leu
 195 200 205
 Gln Met His Val Ser Ile Lys Asn Pro Pro Asn Pro Glu Lys Val Ala
 210 215 220
 Ala Asn Ile Lys Arg Leu Asn Asp Leu Gly Leu Glu Val His Ile Thr
 225 230 235 240
 Glu Met Asp Val Lys Thr Trp Asp Gly Ile Gly Thr Lys Gln Gln Arg
 245 250 255
 Leu Ala Ala Gln Ala Gln Val Tyr Arg Asn Met Met Gln Val Cys Leu
 260 265 270
 Glu Ala Glu Asn Cys Lys Ala Phe Ser Leu Trp Gly Val Ser Asp Arg
 275 280 285
 Tyr Ser Trp Ile Pro Arg Ile Phe Lys Lys Pro Asp Ala Pro Leu Ile
 290 295 300
 Phe Asp Asp Leu Gly Arg Pro Lys Pro Ala Tyr Asn Ala Leu Lys Glu
 305 310 315 320
 Val Leu Lys Arg Arg

325

<210> 27
 <211> 1173
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 27
 atgaaatcct taacaaatca atccttcatg aaactcataa tctgtctggc attgccagtc 60
 gcactactca gcatttcatg caaaaaaccc gccgaaccac tgaaacgggt tgaaggctta 120
 aaagacagct tcaaagacaa gtcttctcatg ggtgtggcgc tgaataaagc acagattctg 180
 ggaagagata cattggtaca tgcttttaca gtacagcatt ttaattccat tactgcagaa 240
 aacgaaatga agtgggaacg catccaccgc cagcctgatg tatatgattt cacgggtccg 300
 gacagcctga ttgcttttgg cgaacgcaac ggcattgtta tagtcgggca tacactcgta 360
 tggcactccc aggtgcccga ttgggttttc accgatgaga agggaaagcc tctgacccgc 420
 gatgctctgc tccaacgcat gaaggatcat atttatgccg ttgtcggccg gtataagggc 480
 aaggtggatg gctgggatgt ggtaaatgaa gcattggatg aagacggaca gctgcgcaaa 540
 tcagggtggc atgaaatcat cgggtgatgat tacattcaga aagcctttga gttcaccgcg 600
 gaggcagatc ccggtgcaga gcctttattac aatgattaca acatagaact caaaaaaag 660
 cgggaggggtg ctgtcaggct gctacaggaa ctgcagcaaa aaggcattaa aatcgacgga 720
 gtgggcattc agggacattg gcacctgcac tcacctgatc tgcaagagat tgattcaagt 780
 cttcaggcat acggacaact ttggtctgaag gtcattgatca ccgaactgga tgtaacgtc 840
 attcccgaac cttcagggtat tattggcgcc gatgttgac agcgggcgga ttatcagagc 900
 cagctgaatc catggcctga aagttttccc gattccatgc agcaggttct ggccagcccg 960
 tatgccgaac tgttcggatt gtctctgaag cacagcgata aggtaagccg ggtgaccttc 1020
 tggggaattc acgatggcta ttcttggaag aacaactggc caataccggg ccgaacaact 1080
 tatcccctcc tttttgaccg gaattaccag cctaaacctg cgtatgatgc tgcattgaa 1140
 ttgacaaaaa tacagccgga agccagtaac tga 1173

<210> 28
 <211> 390
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(27)

<400> 28
 Met Lys Ser Leu Thr Asn Gln Ser Phe Met Lys Leu Ile Ile Cys Leu
 1 5 10 15
 Ala Leu Pro Val Ala Leu Leu Ser Ile Ser Cys Lys Lys Pro Ala Glu
 20 25 30
 Pro Leu Lys Pro Val Glu Gly Leu Lys Asp Ser Phe Lys Asp Lys Phe
 35 40 45
 Leu Met Gly Val Ala Leu Asn Lys Ala Gln Ile Leu Gly Arg Asp Thr
 50 55 60
 Leu Val His Ala Phe Thr Val Gln His Phe Asn Ser Ile Thr Ala Glu
 65 70 75 80
 Asn Glu Met Lys Trp Glu Arg Ile His Pro Gln Pro Asp Val Tyr Asp
 85 90 95
 Phe Thr Val Pro Asp Ser Leu Ile Ala Phe Gly Glu Arg Asn Gly Met
 100 105 110
 Phe Ile Val Gly His Thr Leu Val Trp His Ser Gln Val Pro Asp Trp
 115 120 125
 Val Phe Thr Asp Glu Lys Gly Lys Pro Leu Thr Arg Asp Ala Leu Leu
 130 135 140
 Gln Arg Met Lys Asp His Ile Tyr Ala Val Val Gly Arg Tyr Lys Gly
 145 150 155 160
 Lys Val Asp Gly Trp Asp Val Val Asn Glu Ala Leu Asp Glu Asp Gly
 165 170 175
 Gln Leu Arg Lys Ser Arg Trp His Glu Ile Ile Gly Asp Asp Tyr Ile
 180 185 190
 Gln Lys Ala Phe Glu Phe Thr Arg Glu Ala Asp Pro Gly Ala Glu Leu
 Page 20

195 200 205
 Tyr Tyr Asn Asp Tyr Asn Ile Glu Leu Lys Lys Lys Arg Glu Gly Ala
 210 215 220
 Val Arg Leu Leu Gln Glu Leu Gln Gln Lys Gly Ile Lys Ile Asp Gly
 225 230 235 240
 Val Gly Ile Gln Gly His Trp His Leu His Ser Pro Asp Leu Gln Glu
 245 250 255
 Ile Asp Ser Ser Leu Gln Ala Tyr Gly Gln Leu Gly Leu Lys Val Met
 260 265 270
 Ile Thr Glu Leu Asp Val Asn Val Ile Pro Glu Pro Ser Gly Ile Ile
 275 280 285
 Gly Ala Asp Val Ala Gln Arg Ala Asp Tyr Gln Ser Gln Leu Asn Pro
 290 295 300
 Trp Pro Glu Ser Phe Pro Asp Ser Met Gln Gln Val Leu Ala Ser Arg
 305 310 315 320
 Tyr Ala Glu Leu Phe Gly Leu Phe Leu Lys His Ser Asp Lys Val Ser
 325 330 335
 Arg Val Thr Phe Trp Gly Ile His Asp Gly Tyr Ser Trp Lys Asn Asn
 340 345 350
 Trp Pro Ile Pro Gly Arg Thr Thr Tyr Pro Leu Leu Phe Asp Arg Asn
 355 360 365
 Tyr Gln Pro Lys Pro Ala Tyr Asp Ala Val Ile Glu Leu Thr Lys Ile
 370 375 380
 Gln Pro Glu Ala Ser Asn
 385 390

<210> 29
 <211> 2331
 <212> DNA
 <213> Archaea

<400> 29
 atgacgatgc agagaaagta ctcatccgac gcgaacacac agtatgagtg gataaaatca 60
 gctactgtac catctggtca gtgggtacag ctctctggaa cgtacacgat cccggccgga 120
 gttaccgtgg aagatctcac gctttacttc gaatctcaaa atccaaccct tgagttctac 180
 gtggatgacg tgaagatagt ggatacaact tccgcagaga taaagattga aatggaacct 240
 gaaaaagaga tacctgctct gaaagaagta ctgaaagatt acttcaaagt cggagttgca 300
 ctgccgttga aggtcttctt caaccggaag gacatagaac tcatacagaa acacttcaac 360
 agcatcaccg cagaaaacga gatgaaaccg gatagtctgc tcgcgggcat cgaaaacggt 420
 aagctgaagt tcaggtttga aacagcagac aaatacattc agttcgtcga ggaaaacggc 480
 atggttataa gaggtcacac actggtgtgg cacaaccaga cccccgactg gttcttcaaa 540
 gacgaaaacg gaaacctctt ctcaaagaa gcgatgacgg aaagactcaa agagtacatc 600
 cacaccgttg tcggacactt caaaggaaaa gtctacgcat gggacgtggg gaacgaagcg 660
 gtcgatccga accagccgga tggactgaga agatcaacct ggtaccagat catggggcct 720
 gactacatag aactcgcctt caagttcgca agagaggcag atccagatgc aaaactcttc 780
 tacaacgact acaacacatt cgatcccaga aagagagaca tcacttcaa cctcgtgaag 840
 gatctcaaag agaagggact catcgatggc ataggaatgc agtgtcacat cagtcttgca 900
 acagacatca aacagatcga agaggccatc aaaaagttca gcaccatacc cggatataga 960
 attcacatca cagaactcga tatgagtgtc tacagagatt ccagttccaa ctaccagag 1020
 gcaccgagga cggcactcat cgaacagggt cacaataatga tgcagctctt tgagatcttc 1080
 aagaagcaca gcaacgtgat cacgaacgtc acattctggg gtctcaagga cgattacttc 1140
 tggagagcaa caagaagaaa cgactggccg ctcatcttcg acaaagatca ccaggcgaaa 1200
 ctgccttact gggcgatagt ggcacctgag gtccttcac cacttcaaaa agaaagcagg 1260
 atctccgaag gcgaagcagt ggtagtgggg atgatggacg actcgtacct gatgtcgaag 1320
 ccgatagaga tccttgacga agaaggggaa gtgaaggcaa cgatcagggc agtgtggaaa 1380
 gacagcaga tctatctca cggagagga caggacaaga caaagaaacc agcagaagac 1440
 ggagtggcca tattactcaa cccgaacaac cctatctgca gcctgatgac 1500
 acctacgttg tgctgtggac gaactggaag acggaggtca acagagaaga cgtacagggt 1560
 aagaaattcg ttgggcctgg cttagaaga tacagcttcg agatgtcgat cagcataccg 1620
 ggtgtggagt tcaagaaaga cagctacata ggatttgacg ttgcggtgat agacgacggg 1680
 aagtgttaca gctggagcga cacgacgaac agccagaaga cgaacacgat gaactacgga 1740
 acgctgaagc tcgaaggaat aatggtagcg acagcaaat acggaacacc ggtcatcgat 1800
 ggagagatcg atgagatctg gaacacgaca gaggagatag agacgaaagc ggtggctatg 1860
 ggatcgcttg acaagaatgc gacagcgaaa gtgaggggtc tgtgggacga gaactacctg 1920
 tacgtacttg cgatcgtgaa agagcccgtt tgaacaaaag acaacagcaa cccgtgggag 1980
 caggattccg tggagatctt cgtggatgag aacaaccaca agacaggata ctacgaagac 2040
 gacgacgcgc agttcagggt gaactacatg aacgagcaga cctttggaac gggaggaagt 2100
 ccagcgaggt tcaagacagc ggtgaagctg atcgaaggag gatacatagt tgaggcagcg 2160
 atcaagtggg agacgatcaa gccaacaccg aacacagtga taggattcaa catccagggt 2220

aacgatgcga acgagaaagg gcagagggtc ggtatcatct cctggagcga tcccacaaac 2280
 aacagctggc aagatccttc aaagttcggg aacctcagac tcatcaagtg a 2331

<210> 30
 <211> 776
 <212> PRT
 <213> Archaea

<400> 30
 Met Thr Met Gln Arg Lys Tyr Ser Ser Asp Ala Asn Thr Gln Tyr Glu
 1 5 10 15
 Trp Ile Lys Ser Ala Thr Val Pro Ser Gly Gln Trp Val Gln Leu Ser
 20 25 30
 Gly Thr Tyr Thr Ile Pro Ala Gly Val Thr Val Glu Asp Leu Thr Leu
 35 40 45
 Tyr Phe Glu Ser Gln Asn Pro Thr Leu Glu Phe Tyr Val Asp Asp Val
 50 55 60
 Lys Ile Val Asp Thr Thr Ser Ala Glu Ile Lys Ile Glu Met Glu Pro
 65 70 75 80
 Glu Lys Glu Ile Pro Ala Leu Lys Glu Val Leu Lys Asp Tyr Phe Lys
 85 90 95
 Val Gly Val Ala Leu Pro Ser Lys Val Phe Leu Asn Pro Lys Asp Ile
 100 105 110
 Glu Leu Ile Thr Lys His Phe Asn Ser Ile Thr Ala Glu Asn Glu Met
 115 120 125
 Lys Pro Asp Ser Leu Leu Ala Gly Ile Glu Asn Gly Lys Leu Lys Phe
 130 135 140
 Arg Phe Glu Thr Ala Asp Lys Tyr Ile Gln Phe Val Glu Glu Asn Gly
 145 150 155 160
 Met Val Ile Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro Asp
 165 170 175
 Trp Phe Phe Lys Asp Glu Asn Gly Asn Leu Leu Ser Lys Glu Ala Met
 180 185 190
 Thr Glu Arg Leu Lys Glu Tyr Ile His Thr Val Val Gly His Phe Lys
 195 200 205
 Gly Lys Val Tyr Ala Trp Asp Val Val Asn Glu Ala Val Asp Pro Asn
 210 215 220
 Gln Pro Asp Gly Leu Arg Arg Ser Thr Trp Tyr Gln Ile Met Gly Pro
 225 230 235 240
 Asp Tyr Ile Glu Leu Ala Phe Lys Phe Ala Arg Glu Ala Asp Pro Asp
 245 250 255
 Ala Lys Leu Phe Tyr Asn Asp Tyr Asn Thr Phe Asp Pro Arg Lys Arg
 260 265 270
 Asp Ile Ile Tyr Asn Leu Val Lys Asp Leu Lys Glu Lys Gly Leu Ile
 275 280 285
 Asp Gly Ile Gly Met Gln Cys His Ile Ser Leu Ala Thr Asp Ile Lys
 290 295 300
 Gln Ile Glu Glu Ala Ile Lys Lys Phe Ser Thr Ile Pro Gly Ile Glu
 305 310 315 320
 Ile His Ile Thr Glu Leu Asp Met Ser Val Tyr Arg Asp Ser Ser Ser
 325 330 335
 Asn Tyr Pro Glu Ala Pro Arg Thr Ala Leu Ile Glu Gln Ala His Lys
 340 345 350
 Met Met Gln Leu Phe Glu Ile Phe Lys Lys His Ser Asn Val Ile Thr
 355 360 365
 Asn Val Thr Phe Trp Gly Leu Lys Asp Asp Tyr Ser Trp Arg Ala Thr
 370 375 380
 Arg Arg Asn Asp Trp Pro Leu Ile Phe Asp Lys Asp His Gln Ala Lys
 385 390 395 400
 Leu Ala Tyr Trp Ala Ile Val Ala Pro Glu Val Leu Pro Pro Leu Pro
 405 410 415
 Lys Glu Ser Arg Ile Ser Glu Gly Glu Ala Val Val Val Gly Met Met
 420 425 430
 Asp Asp Ser Tyr Leu Met Ser Lys Pro Ile Glu Ile Leu Asp Glu Glu
 435 440 445
 Gly Asn Val Lys Ala Thr Ile Arg Ala Val Trp Lys Asp Ser Thr Ile
 450 455 460
 Tyr Ile Tyr Gly Glu Val Gln Asp Lys Thr Lys Lys Pro Ala Glu Asp
 465 470 475 480

Gly Val Ala Ile Phe Ile Asn Pro Asn Asn Glu Arg Thr Pro Tyr Leu
 485 490 495
 Gln Pro Asp Asp Thr Tyr Val Val Leu Trp Thr Asn Trp Lys Thr Glu
 500 505 510
 Val Asn Arg Glu Asp Val Gln Val Lys Lys Phe Val Gly Pro Gly Phe
 515 520 525
 Arg Arg Tyr Ser Phe Glu Met Ser Ile Thr Ile Pro Gly Val Glu Phe
 530 535 540
 Lys Lys Asp Ser Tyr Ile Gly Phe Asp Val Ala Val Ile Asp Asp Gly
 545 550 555 560
 Lys Trp Tyr Ser Trp Ser Asp Thr Thr Asn Ser Gln Lys Thr Asn Thr
 565 570 575
 Met Asn Tyr Gly Thr Leu Lys Leu Glu Gly Ile Met Val Ala Thr Ala
 580 585 590
 Lys Tyr Gly Thr Pro Val Ile Asp Gly Glu Ile Asp Glu Ile Trp Asn
 595 600 605
 Thr Thr Glu Glu Ile Glu Thr Lys Ala Val Ala Met Gly Ser Leu Asp
 610 615 620
 Lys Asn Ala Thr Ala Lys Val Arg Val Leu Trp Asp Glu Asn Tyr Leu
 625 630 635 640
 Tyr Val Leu Ala Ile Val Lys Glu Pro Val Leu Asn Lys Asp Asn Ser
 645 650 655
 Asn Pro Trp Glu Gln Asp Ser Val Glu Ile Phe Val Asp Glu Asn Asn
 660 665 670
 His Lys Thr Gly Tyr Tyr Glu Asp Asp Ala Gln Phe Arg Val Asn
 675 680 685
 Tyr Met Asn Glu Gln Thr Phe Gly Thr Gly Gly Ser Pro Ala Arg Phe
 690 695 700
 Lys Thr Ala Val Lys Leu Ile Glu Gly Gly Tyr Ile Val Glu Ala Ala
 705 710 715 720
 Ile Lys Trp Lys Thr Ile Lys Pro Thr Pro Asn Thr Val Ile Gly Phe
 725 730 735
 Asn Ile Gln Val Asn Asp Ala Asn Glu Lys Gly Gln Arg Val Gly Ile
 740 745 750
 Ile Ser Trp Ser Asp Pro Thr Asn Asn Ser Trp Gln Asp Pro Ser Lys
 755 760 765
 Phe Gly Asn Leu Arg Leu Ile Lys
 770 775

<210> 31
 <211> 1134
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 31
 gtggaaaccg tcggagcacc ggagctgagc tatgaaatcc ggaatttccg ggtggtggca 60
 ccggacggag tgccggatat acagcccaca gccgcaccgg aagcgcaggc tgttccggaa 120
 ggggagatgc cttccctgaa ggatgtatac gcgggcaaat tcgacttcgg tacggcgctg 180
 ccccggaatg cattcaatga tatccagctg ctgagactgg tgaaggacca gttcaacatc 240
 ctgacaccgg aaaatgagat gaaaccggat gcaatcctgg atgtgtacgg cagcaaaaaa 300
 ctggcggaag aggacgagac agcgggtggct gtccggtttg aagcatgcaa gacgctgctt 360
 cggttcgcac agtccaacgg cctgaagggtg cacggccata cgctgctgtg gcacaaccag 420
 accccggaag cccctttcca cgaaggttat gacaccacca agccgatggc cggccgggaa 480
 gtgatgttgg gccggatgga gaattacatc cgcgaaagtgc tgacctggac cgaagaaaat 540
 tatccgggag tgatcgtttc ctgggacgtg gtgaatgaag caatcgacga cggaacgaac 600
 cagctgcgca ccggtgccaa ctggtataag acggtcggac cggactacct ggcacgcgcg 660
 tttgaatatg cccggaaata cgcggcggaa ggcgtgctgc tgtactacaa cgattacaat 720
 accgcatacg gcggtaagct gtatgggatt gtggatctgc tggagagcct gattgccgag 780
 ggcaatattg acggatacgg attccagatg caccacagcc tgggagaacc ttccatggat 840
 atgattaccc gggcagtaga gaaaatagcc tcgctgggac tccggctgcg tgtgagcgaa 900
 ctggacatca acgcccggaa ggcgacagag aaaaatttcg aagcccagaa gaacaagtac 960
 aaacaggtga tgaagctgat gctccggttc aaggaccaga ctgaagcggc ccagggtgtg 1020
 ggcgtagcgg acatcatgag ctggcgagg gacggatatc cgctgctgtt tgacaagaac 1080
 atgaatccga aaccgcggtt ctccggtgtg atcgaagccg gaatggaaga ctga 1134

<210> 32

<211> 377
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 32
 Val Glu Thr Val Gly Ala Pro Glu Leu Ser Tyr Glu Ile Arg Asn Phe
 1 5 10 15
 Arg Val Val Ala Pro Asp Gly Val Pro Asp Ile Gln Pro Thr Ala Ala
 20 25 30
 Pro Glu Ala Gln Ala Val Pro Glu Gly Glu Met Pro Ser Leu Lys Asp
 35 40 45
 Val Tyr Ala Gly Lys Phe Asp Phe Gly Thr Ala Leu Pro Arg Asn Ala
 50 55 60
 Phe Asn Asp Ile Gln Leu Leu Arg Leu Val Lys Asp Gln Phe Asn Ile
 65 70 75 80
 Leu Thr Pro Glu Asn Glu Met Lys Pro Asp Ala Ile Leu Asp Val Tyr
 85 90 95
 Gly Ser Lys Lys Leu Ala Glu Lys Asp Glu Thr Ala Val Ala Val Arg
 100 105 110
 Phe Glu Ala Cys Lys Thr Leu Leu Arg Phe Ala Gln Ser Asn Gly Leu
 115 120 125
 Lys Val His Gly His Thr Leu Leu Trp His Asn Gln Thr Pro Glu Ala
 130 135 140
 Leu Phe His Glu Gly Tyr Asp Thr Thr Lys Pro Met Ala Gly Arg Glu
 145 150 155 160
 Val Met Leu Gly Arg Met Glu Asn Tyr Ile Arg Glu Val Leu Thr Trp
 165 170 175
 Thr Glu Glu Asn Tyr Pro Gly Val Ile Val Ser Trp Asp Val Val Asn
 180 185 190
 Glu Ala Ile Asp Asp Gly Thr Asn Gln Leu Arg Thr Gly Ala Asn Trp
 195 200 205
 Tyr Lys Thr Val Gly Pro Asp Tyr Leu Ala Arg Ala Phe Glu Tyr Ala
 210 215 220
 Arg Lys Tyr Ala Ala Glu Gly Val Leu Leu Tyr Tyr Asn Asp Tyr Asn
 225 230 235 240
 Thr Ala Tyr Gly Gly Lys Leu Tyr Gly Ile Val Asp Leu Leu Glu Ser
 245 250 255
 Leu Ile Ala Glu Gly Asn Ile Asp Gly Tyr Gly Phe Gln Met His His
 260 265 270
 Ser Leu Gly Glu Pro Ser Met Asp Met Ile Thr Arg Ala Val Glu Lys
 275 280 285
 Ile Ala Ser Leu Gly Leu Arg Leu Arg Val Ser Glu Leu Asp Ile Asn
 290 295 300
 Ala Gly Lys Ala Thr Glu Lys Asn Phe Glu Ala Gln Lys Asn Lys Tyr
 305 310 315 320
 Lys Gln Val Met Lys Leu Met Leu Arg Phe Lys Asp Gln Thr Glu Ala
 325 330 335
 Val Gln Val Trp Gly Val Thr Asp Ile Met Ser Trp Arg Arg Asp Gly
 340 345 350
 Tyr Pro Leu Leu Phe Asp Lys Asn Met Asn Pro Lys Pro Ala Phe Phe
 355 360 365
 Gly Val Ile Glu Ala Gly Met Glu Asp
 370 375

<210> 33
 <211> 1815
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 33
 atggttcgca aaaaactatt ttatatcgctc gcgttaatgc tgatgttcgg cgcaagtttt
 acttcgctc aggacgcgga attttccctg cgcggttag ccgagcgcaa taacttttat
 gttggagcag ccgtttatac cactcatctg aatgatcctg tccatgttga aacactggca

60
 120
 180

cgagaattca	atatgctcac	gcctgaacag	caggccaaac	attgtgagtt	ggaggcacag	240
caaggtcaat	ttgactttcg	gagtttcgat	cgttttagtcg	ccttcgccga	agaacacaac	300
atggcgatac	acgggtcatgc	gctgggtctgg	catagctgca	caccgcaatg	ggtggctaac	360
ggcgagtaca	cccgtgacga	agccatttgg	ctgctgcgcg	actcgattat	gaccattgtt	420
ggccgttaca	aaggccgtat	tccgatttgg	gacgtcgtca	atgaaggcat	tgctgacagc	480
ggcggaacac	tgcgcgatac	gccatggcgg	cagtttaattg	gcgatgatta	catcgaactt	540
gccttccagt	tcgctcatga	agccgacccg	gatgcgctgc	tgttttacaa	cgactataat	600
acggaaggca	tgaaccctaa	atcggacgcc	atgtacgaga	tggtgagcga	ttttgtggcg	660
cgtggaattc	cgattcacgg	ggttgggctg	caatcccatt	tcatattagg	cagttttgac	720
ccagaccaga	ttgctcggaa	cgtcgcgcgg	cttggcgaa	tcggtttaca	agttcaattc	780
accgaggtcg	atattcgata	ttccggcgag	gcgacagata	atatacctcca	gcggcgaggcg	840
ggcgattacc	atcgcctgat	ggacgtttgc	ctcggtaacg	acgcctgtac	tgcgtttatc	900
acctggggcg	tgaccgataa	atatacctgg	ttgcggggcg	cgaacctggg	cttctacaac	960
aacctatcgg	ttgaaccgct	gctttttgac	gatgactatg	aaccaagcc	cgcttatttt	1020
gcggtgctgg	actcactagc	gcgaagagcg	ggcgaaaccc	ccgttttgag	cgatgacgag	1080
cttgcggcga	tgatcggcgg	cacagtccaa	acggtcgaaa	ttccccgcc	gacgaaaagc	1140
aatctcagtc	aggaagcgcc	ggacgccgtt	cctggtgtga	tctattacgc	cgctacccc	1200
ataagcatca	cagttgacgg	cgaagccaac	gattgggaac	gcattccgcg	cggtatgatt	1260
gacagcggcc	ccaccgtacc	tcaggataac	gacacgacaa	tgacatttgc	cgccgctgcc	1320
gacaaaacca	atctatactt	ccttgacagag	gttacggaca	gccagggtgtc	ctacgggaacg	1380
cacgaccggg	ctactgcctg	gtatcaggag	gactcgggtg	agttttacct	gaacacgaca	1440
ggcgatctaa	ctaaccaccg	ctaccaaccc	ggcgtcgcgc	aaatcggtat	catggcagcc	1500
aacatcgaca	acgataatcc	cgggtgcaccg	atcatcgggg	gcggcaacag	cgacatttcg	1560
caggtaaaag	cgattgtcgt	caaaaccgat	accgggtatc	tggtcgaggc	gtctgttcca	1620
ctcatgaccg	atgtctggac	gattgaaccg	aaacaagggg	ctgtactcgg	cttccaagtg	1680
catctcaatg	gctcacgcac	accggatgcc	gaccgagaca	ccaagttgat	ctggtcgccta	1740
ctggatacgc	tagatcagtc	ctatagcaat	cccagcctgt	ttggccgact	catcttctgg	1800
aacataaatc	tctaa					1815

<210> 34
 <211> 604
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(23)

<400> 34
 Met Val Arg Lys Lys Leu Phe Tyr Ile Val Ala Leu Met Leu Met Phe
 1 5 10 15
 Gly Ala Ser Phe Thr Ser Ala Gln Asp Ala Glu Phe Ser Leu Arg Gly
 20 25 30
 Leu Ala Glu Arg Asn Asn Phe Tyr Val Gly Ala Ala Val Tyr Thr Thr
 35 40 45
 His Leu Asn Asp Pro Val His Val Glu Thr Leu Ala Arg Glu Phe Asn
 50 55 60
 Met Leu Thr Pro Glu Gln Gln Ala Lys His Cys Glu Leu Glu Ala Gln
 65 70 75 80
 Gln Gly Gln Phe Asp Phe Arg Ser Phe Asp Arg Leu Val Ala Phe Ala
 85 90 95
 Glu Glu His Asn Met Ala Ile His Gly His Ala Leu Val Trp His Ser
 100 105 110
 Cys Thr Pro Gln Trp Val Ala Asn Gly Glu Tyr Thr Arg Asp Glu Ala
 115 120 125
 Ile Gly Leu Leu Arg Asp Ser Ile Met Thr Ile Val Gly Arg Tyr Lys
 130 135 140
 Gly Arg Ile Pro Ile Trp Asp Val Val Asn Glu Gly Ile Ala Asp Ser
 145 150 155 160
 Gly Gly Thr Leu Arg Asp Thr Pro Trp Arg Gln Leu Ile Gly Asp Asp
 165 170 175
 Tyr Ile Glu Leu Ala Phe Gln Phe Ala His Glu Ala Asp Pro Asp Ala
 180 185 190
 Leu Leu Phe Tyr Asn Asp Tyr Asn Thr Glu Gly Met Asn Pro Lys Ser
 195 200 205
 Asp Ala Met Tyr Glu Met Val Ser Asp Phe Val Ala Arg Gly Ile Pro
 210 215 220

<210>	35
<211>	2286
<212>	DNA
<213>	Unknown

<220>
<223> obtained from an environmental sample

<400>	35						
atgaccttga	ttacgccaa	ctcgaaatta	accctcacta	aagggaacaa	aagctggagc		60
tcgcgcgcct	gcaggctgac	actagtggat	ctcacgcttt	acttcgaatc	tcaaaatcca		120
acccttgagt	tctacgtgga	tgacgtgaa	atagtggata	caacttccgc	agagataaa		180
attgaaattg	aaactgaaaa	agagataacct	gctctgaaag	aagtactgaa	agattacttc		240
aaagctcggag	ttgcactgcc	gtccaagggtc	ttcttcaacc	cgaaggacat	agaactcatc		300
acgaaacact	tcaacagcat	caccgcagaa	aacgagatga	aaccggatag	tctgctcgcg		360
ggcatcgaaa	acggttaagct	gaagttcagg	tttgaacag	cagacaaata	cattcagttc		420
gtcgaggaaa	acggcatggt	tataagaggt	cacacactgg	tgtggcacia	ccagacaccc		480
gactggttct	tcaaagacga	aaacggaaac	ctcctctcca	aagaagcgat	gcggaaaga		540
ctcaaagagt	acatccacac	cgttgctcga	cacttcaaag	gaaaagtcta	cgcattgggac		600
gtggtgaaag	aagcggtcga	tccgaaccag	ccggtatggac	tgagaagatc	aacctgggtac		660

cgatcatg	ggcctgacta	catagaactc	gccttcaagt	tcgcaagaga	ggcagatcca	720
gatgcaaaac	tctttctacaa	cgactacaac	acattcgatc	ccagaaagag	agacatcatc	780
tacaacctcg	tgaaggatct	caaagagaag	ggactcatcg	atggcatagg	aatgcagtgt	840
cacatcagtc	ttgcaacaga	catcaaacag	atcgaagagg	ccatcaaaaa	gttcagcacc	900
atacccggta	tagaaattca	catcacagaa	ctcgatatga	gtgtctacag	agattccagt	960
tccaactacc	cagaggcacc	gaggacggca	ctcatcgaac	aggctcacia	aatgatgcag	1020
ctctttgaga	tcttcaagaa	gcacagcaac	gtgatcacga	acgtcacatt	ctggggtctc	1080
aaggacgatt	actcctggag	agcaacaaga	agaaacgact	ggccgctcat	cttcgacaaa	1140
gatcaccagg	cgaaactcgc	ttactgggcy	atagtggcac	ctgaggctct	tccaccactt	1200
ccaaaagaaa	gcaggatctc	cgaaggcgaa	gcagtggtag	tggggatgat	ggacgactcg	1260
tacctgatgt	cgaagccgat	agagatcctt	gacgaagaag	ggaacgtgaa	ggcaacgata	1320
agggcagtgt	ggaaagacag	cacgatctac	atctacggag	aggtagagga	caagacaaag	1380
aaaccagcag	aagacggagt	ggccatattc	atcaacccga	acaacgaaag	aacaccctat	1440
ctgcagcctg	atgacaccta	cgttgtgctg	tggacgaact	ggaagacgga	ggtcaacaga	1500
gaagacgtac	aggtgaagaa	attcgttggg	cctggcttta	gaagatacag	cttcgagatg	1560
tcgatcacga	taccgggtgt	ggagttcaag	aaagacagct	acataggatt	tgacgttgcy	1620
gtgatagacg	acgggaagtgt	gtacagctgg	agcgacacga	cgaacagcca	gaagacgaac	1680
acgatgaact	acggaacgct	gaagctcgaa	ggaataatgg	tagcgacagc	aaaatacggg	1740
acaccgggtca	tcgatggaga	gatcgatgag	atctggaaca	cgacagagga	gatagagacg	1800
aaagcgggtgg	ctatgggagt	gcttgacaag	aatgcgacag	cgaaagtggg	ggtgctgtgg	1860
gacgagaact	actgtgacgt	acttgcgac	gtgaaagatc	ccgttctgaa	caaagacaa	1920
agcaaccctg	gggagcagga	ttccgtggag	atcttctgtg	atgagaacaa	ccacaagaca	1980
ggatactacg	aagacgacga	cgcgacgttc	agggtgaact	acatgaacga	gcagaccttt	2040
ggaacgggag	gaagtcacgc	gaggttcaag	acagcgggtg	agctgatcga	aggaggatag	2100
atagttgagc	cagcgatcaa	gtggaagacg	atcaagccaa	caccgaacac	agtgatagga	2160
ttcaacatcc	aggtagaacga	tgcaaacgag	aaagggcaga	gggtcggtat	catctcctgg	2220
agcgatccca	caaacaacag	ctggcaagat	ccttcaaagt	tcggtaacct	cagactcatc	2280
aagtga						2286

<210> 36
 <211> 761
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 36

Met	Thr	Leu	Ile	Thr	Pro	Ser	Ser	Lys	Leu	Thr	Leu	Thr	Lys	Gly	Asn
1				5				10					15		
Lys	Ser	Trp	Ser	Ser	Arg	Ala	Cys	Arg	Ser	Thr	Leu	Val	Asp	Leu	Thr
		20						25					30		
Leu	Tyr	Phe	Glu	Ser	Gln	Asn	Pro	Thr	Leu	Glu	Phe	Tyr	Val	Asp	Asp
		35					40					45			
Val	Lys	Ile	Val	Asp	Thr	Thr	Ser	Ala	Glu	Ile	Lys	Ile	Glu	Met	Glu
	50				55					60					
Pro	Glu	Lys	Glu	Ile	Pro	Ala	Leu	Lys	Glu	Val	Leu	Lys	Asp	Tyr	Phe
65					70				75					80	
Lys	Val	Gly	Val	Ala	Leu	Pro	Ser	Lys	Val	Phe	Leu	Asn	Pro	Lys	Asp
		85							90					95	
Ile	Glu	Leu	Ile	Thr	Lys	His	Phe	Asn	Ser	Ile	Thr	Ala	Glu	Asn	Glu
		100						105					110		
Met	Lys	Pro	Asp	Ser	Leu	Leu	Ala	Gly	Ile	Glu	Asn	Gly	Lys	Leu	Lys
		115					120					125			
Phe	Arg	Phe	Glu	Thr	Ala	Asp	Lys	Tyr	Ile	Gln	Phe	Val	Glu	Glu	Asn
	130					135					140				
Gly	Met	Val	Ile	Arg	Gly	His	Thr	Leu	Val	Trp	His	Asn	Gln	Thr	Pro
145					150				155						160
Asp	Trp	Phe	Phe	Lys	Asp	Glu	Asn	Gly	Asn	Leu	Leu	Ser	Lys	Glu	Ala
		165						170						175	
Met	Thr	Glu	Arg	Leu	Lys	Glu	Tyr	Ile	His	Thr	Val	Val	Gly	His	Phe
		180						185					190		
Lys	Gly	Lys	Val	Tyr	Ala	Trp	Asp	Val	Val	Asn	Glu	Ala	Val	Asp	Pro
		195					200					205			
Asn	Gln	Pro	Asp	Gly	Leu	Arg	Arg	Ser	Thr	Trp	Tyr	Gln	Ile	Met	Gly
	210					215					220				
Pro	Asp	Tyr	Ile	Glu	Leu	Ala	Phe	Lys	Phe	Ala	Arg	Glu	Ala	Asp	Pro
225					230					235					240
Asp	Ala	Lys	Leu	Phe	Tyr	Asn	Asp	Tyr	Asn	Thr	Phe	Asp	Pro	Arg	Lys

Arg Asp Ile Ile Tyr Asn Leu Val Lys Asp Leu Lys Glu Lys Gly Leu
 Ile Asp Gly Ile Gly Met Gln Cys His Ile Ser Leu Ala Thr Asp Ile
 Lys Gln Ile Glu Glu Ala Ile Lys Lys Phe Ser Thr Ile Pro Gly Ile
 Glu Ile His Ile Thr Glu Leu Asp Met Ser Val Tyr Arg Asp Ser Ser
 Ser Asn Tyr Pro Glu Ala Pro Arg Thr Ala Leu Ile Glu Gln Ala His
 Lys Met Met Gln Leu Phe Glu Ile Phe Lys Lys His Ser Asn Val Ile
 Thr Asn Val Thr Phe Trp Gly Leu Lys Asp Asp Tyr Ser Trp Arg Ala
 Thr Arg Arg Asn Asp Trp Pro Leu Ile Phe Asp Lys Asp His Gln Ala
 Lys Leu Ala Tyr Trp Ala Ile Val Ala Pro Glu Val Leu Pro Pro Leu
 Pro Lys Glu Ser Arg Ile Ser Glu Gly Glu Ala Val Val Val Gly Met
 Met Asp Asp Ser Tyr Leu Met Ser Lys Pro Ile Glu Ile Leu Asp Glu
 Glu Gly Asn Val Lys Ala Thr Ile Arg Ala Val Trp Lys Asp Ser Thr
 Ile Tyr Ile Tyr Gly Glu Val Gln Asp Lys Thr Lys Lys Pro Ala Glu
 Asp Gly Val Ala Ile Phe Ile Asn Pro Asn Asn Glu Arg Thr Pro Tyr
 Leu Gln Pro Asp Asp Thr Tyr Val Val Leu Trp Thr Asn Trp Lys Thr
 Glu Val Asn Arg Glu Asp Val Gln Val Lys Lys Phe Val Gly Pro Gly
 Phe Arg Arg Tyr Ser Phe Glu Met Ser Ile Thr Ile Pro Gly Val Glu
 Phe Lys Lys Asp Ser Tyr Ile Gly Phe Asp Val Ala Val Ile Asp Asp
 Gly Lys Trp Tyr Ser Trp Ser Asp Thr Thr Asn Ser Gln Lys Thr Asn
 Thr Met Asn Tyr Gly Thr Leu Lys Leu Glu Gly Ile Met Val Ala Thr
 Ala Lys Tyr Gly Thr Pro Val Ile Asp Gly Glu Ile Asp Glu Ile Trp
 Asn Thr Thr Glu Glu Ile Glu Thr Lys Ala Val Ala Met Gly Ser Leu
 Asp Lys Asn Ala Thr Ala Lys Val Arg Val Leu Trp Asp Glu Asn Tyr
 Leu Tyr Val Leu Ala Ile Val Lys Asp Pro Val Leu Asn Lys Asp Asn
 Ser Asn Pro Trp Glu Gln Asp Ser Val Glu Ile Phe Val Asp Glu Asn
 Asn His Lys Thr Gly Tyr Tyr Glu Asp Asp Ala Gln Phe Arg Val
 Asn Tyr Met Asn Glu Gln Thr Phe Gly Thr Gly Gly Ser Pro Ala Arg
 Phe Lys Thr Ala Val Lys Leu Ile Glu Gly Gly Tyr Ile Val Glu Ala
 Ala Ile Lys Trp Lys Thr Ile Lys Pro Thr Pro Asn Thr Val Ile Gly
 Phe Asn Ile Gln Val Asn Asp Ala Asn Glu Lys Gly Gln Arg Val Gly
 Ile Ile Ser Trp Ser Asp Pro Thr Asn Ser Trp Gln Asp Pro Ser
 Lys Phe Gly Asn Leu Arg Leu Ile Lys

<210> 37
 <211> 2769
 <212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 37

atgcacaaga	agaacggcac	ctattacctg	agctactcta	caaatccggc	caacggcatg	60
cggattgact	acatgaccag	cacgagcccg	accagtggct	ttgtccatcg	aggcacggtc	120
atggcgagc	cctggcagaa	cagcaacaac	aacaaccacg	caaccagcac	cgagtacaac	180
gggcagggct	acatcttcta	tcacaaccgt	gcgttgctga	acgagcgtgc	gggtggcaac	240
gtgctgcagc	gctcggtgaa	cgtggatcgc	ctctacttca	atgccgatgg	cagcatccgt	300
caggtcactt	ccagtgaac	cggcgtgccg	gccctgaaaa	ccctggatgc	cttcctggtc	360
aagcctgccg	agctgtatca	caaggaaagc	gggatcaaga	ccgagcctgc	cagtgaagga	420
accaggcac	tggttatgac	ggctggtagc	tgggtgcgcc	tggccaatgt	cgatttcggc	480
aatggcggcg	ccactggttt	ttccgcgcgt	attgcgga	ccggcagcgg	cagcatccag	540
gtgatcctgg	gcaatctgaa	caacgcccg	gtcggcacgc	tggcagtgg	cagcacccgg	600
aacctccaga	cctggcaaga	ccgcagcacc	gccatcagca	aggtagccgg	cgtgcatgac	660
gtgtatttgc	gtgccaccgg	caatgtgcat	gtgcagcgtc	actggttcgt	ggcgtcggcg	720
ccggccgctg	ccgcctcatc	cagcagtcag	gcaagcgtct	ctgccagcag	tcaggcaagt	780
gtttcttcca	gtagccaggc	aagcgttgcc	tccagcagca	gttccagccg	cgcttcttcc	840
ggcagcagtt	cggtggcggc	tggccaggtg	gaggtcgggt	atcgcccttag	cagcgaatgg	900
gccgcggcgt	tctgcggcgt	ggtcaccatc	cgtaatccgg	gtagtctctc	ggtcaccagc	960
tggagtggca	gtttcaacct	gcctggcggc	aagatcaccc	agctgtggaa	tgccaactgg	1020
accagaaagc	gcagcacctg	gacggtatct	tcccaggcct	ggagcggctg	cattgctgca	1080
ggcgccacca	tcaccacgcc	gggcttctgc	gccgagcgca	cgagcagcaa	tgctctctcc	1140
agtgtcgcca	gcagcagtg	ctcctcatcg	agcagcagtg	ctgcggctgc	cagctccagc	1200
gcggcttcca	gcgtcccgtc	caactggcag	ggcggggctg	gcagcagcgc	atcctcggct	1260
agcagtggcg	ctgcacccaa	gggcgtgctt	gaagtcggcc	tgagcggcct	ttccagccag	1320
gccatgttcc	ccccgttgcg	ggtcaggacg	gacgtgcgg	ccgccaacaa	ggcctatgtt	1380
gaatggccca	acaacggcgc	caatcagtcg	ctggcaacgc	ctgccaacga	tgccgcaggg	1440
caggtggagg	tagccttctg	gctggcccg	gcatccgcag	tgcagtttga	tatcgaagcg	1500
aatttcgcca	acgcggaaga	cgactccttc	tacttccagc	tcaacgggtg	tgccctggcag	1560
accttcaaca	acgccaccac	ggtcggctgg	cagaccctgc	cggtcgcttc	tctgggcaat	1620
ctggctggcg	ggcgccatgt	gctgacctgc	ctgcgcggcg	aggatggcgc	gaagctgggc	1680
aaggtcgctc	tgagtgcggc	acagagcagc	atcagtcgtg	ccacgcccgt	ggcctacgcg	1740
tcgccgaatg	atgttgccaa	cctgttcaag	ctggccagct	tcccgatcgg	ggtggcggtc	1800
agtgcgggca	acgaaggtga	cagcctgctg	cgtagcggta	cccgcgcagc	agccgagcgt	1860
gcgctgaccg	agaagcactt	caacagtcgt	gtggccggca	acatcatgaa	gatgagctac	1920
ctgcataccg	ccgagaacac	ctacaccttc	acccaggcgg	atgcgctggc	cgactacgcc	1980
aagtccaagg	gcatgtgtgt	gcatggccat	gcgctggctt	ggcatgcgga	ctatcaggta	2040
cccaactgga	tgaagaatta	caccggagac	tggctgaaga	tgctcgaagc	ccacgtcacc	2100
accgtcgcca	agcactatgc	cggcaagggt	gtgagctggg	atgtgttgaa	tgaagccctg	2160
gccgatggca	atgccaccgc	caccaagggt	ttccgtgcca	ccgattcgat	cttctatcag	2220
aagatgggct	ccagtctcat	cgagaaggcc	tttattctgc	cacgtgctgc	cgacccgaat	2280
gccgacctgt	attacaacga	ctacggcatg	gagggcggaa	acagcaagtt	caattactgc	2340
atggccatgg	tcgatgattt	ccagaagcgt	ggcattccca	tcgacggcat	cggtttccag	2400
atgcacatca	acatcgactg	gccttcgtcg	gcccgatcc	gcgctgtatt	cagtgaagtg	2460
gtcaagcgtg	gtctgaagg	gcgtatctcc	gagctggata	ttccgggtgaa	taccactgcc	2520
ggctggtttg	ccagcctgaa	tgccacggcc	aacgagctgc	agaagaagaa	gtatcgtgag	2580
gttggtggctg	cctacctgga	tgtggtgccg	cccagctgc	gcggtggcat	caccgtgtgg	2640
ggcctgagcg	acaacggcag	ctggctgggt	acccccacca	agccggactg	gccgctgctg	2700
ttcgtatgcc	acctcaaggc	caaggacgcc	ctgagcggct	ttgccgacgc	cctgcgcggc	2760
gtacgctga						2769

<210> 38

<211> 922

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 38

Met	His	Lys	Lys	Asn	Gly	Thr	Tyr	Tyr	Leu	Ser	Tyr	Ser	Thr	Asn	Pro
1				5					10					15	
Ala	Asn	Gly	Met	Arg	Ile	Asp	Tyr	Met	Thr	Ser	Thr	Ser	Pro	Thr	Ser
			20					25					30		
Gly	Phe	Val	His	Arg	Gly	Thr	Val	Met	Ala	Gln	Pro	Trp	Gln	Asn	Ser
	35						40					45			

Asn	Asn	Asn	Asn	His	Ala	Thr	Ser	Thr	Glu	Tyr	Asn	Gly	Gln	Gly	Tyr
Ile	Phe	Tyr	His	Asn	Arg	Ala	Leu	Ser	Asn	Glu	Arg	Ala	Gly	Gly	Asn
Val	Leu	Gln	Arg	Ser	Val	Asn	Val	Asp	Arg	Leu	Tyr	Phe	Asn	Ala	Asp
Gly	Ser	Ile	Arg	Gln	Val	Thr	Ser	Ser	Ala	Thr	Gly	Val	Pro	Ala	Leu
Lys	Thr	Leu	Asp	Ala	Phe	Leu	Val	Lys	Pro	Ala	Glu	Leu	Tyr	His	Lys
Glu	Ser	Gly	Ile	Lys	Thr	Glu	Pro	Ala	Ser	Glu	Gly	Thr	Gln	Ala	Leu
Val	Met	Thr	Ala	Gly	Ser	Trp	Val	Arg	Leu	Ala	Asn	Val	Asp	Phe	Gly
Asn	Gly	Gly	Ala	Thr	Gly	Phe	Ser	Ala	Arg	Ile	Ala	Ala	Thr	Gly	Ser
Gly	Ser	Ile	Gln	Val	Ile	Leu	Gly	Asn	Leu	Asn	Asn	Ala	Pro	Val	Gly
Thr	Leu	Ala	Val	Ser	Ser	Thr	Gly	Asn	Leu	Gln	Thr	Trp	Gln	Asp	Arg
Ser	Thr	Ala	Ile	Ser	Lys	Val	Thr	Gly	Val	His	Asp	Val	Tyr	Leu	Arg
Ala	Thr	Gly	Asn	Val	His	Val	Gln	Arg	His	Trp	Phe	Val	Ala	Ser	Ala
Pro	Ala	Ala	Ala	Ala	Ser	Ser	Ser	Ser	Gln	Ala	Ser	Val	Ser	Ala	Ser
Ser	Gln	Ala	Ser	Val	Ser	Ser	Ser	Ser	Gln	Ala	Ser	Val	Ala	Ser	Ser
Ser	Ser	Ser	Ser	Arg	Ala	Ser	Ser	Ala	Ser	Ser	Ser	Val	Ala	Ala	Gly
Gln	Val	Glu	Val	Gly	Tyr	Arg	Leu	Ser	Ser	Glu	Trp	Ala	Ala	Gly	Phe
Cys	Gly	Val	Val	Thr	Ile	Arg	Asn	Pro	Gly	Ser	Ser	Pro	Val	Thr	Ser
Trp	Ser	Gly	Ser	Phe	Asn	Leu	Pro	Gly	Gly	Lys	Ile	Thr	Gln	Leu	Trp
Asn	Ala	Asn	Trp	Thr	Gln	Asn	Gly	Ser	Thr	Val	Thr	Val	Ser	Ser	Gln
Ala	Trp	Ser	Gly	Ala	Ile	Ala	Ala	Gly	Ala	Thr	Ile	Thr	Thr	Pro	Gly
Phe	Cys	Ala	Glu	Arg	Thr	Ser	Ser	Asn	Ala	Ser	Ser	Ser	Val	Ala	Ser
Ser	Ser	Val	Ser	Ser	Ser	Ser	Ser	Ser	Ala	Ala	Ala	Ala	Ser	Ser	Ser
Ala	Ala	Ser	Ser	Val	Pro	Ser	Thr	Gly	Ser	Gly	Gly	Val	Gly	Ser	Ser
Ala	Ser	Ser	Ala	Ser	Ser	Ala	Ala	Ala	Pro	Lys	Gly	Val	Leu	Glu	Val
Gly	Leu	Ser	Gly	Leu	Ser	Ser	Gln	Ala	Met	Phe	Ala	Pro	Leu	Arg	Val
Arg	Thr	Asp	Ala	Ala	Ala	Ala	Asn	Lys	Ala	Tyr	Val	Glu	Trp	Pro	Asn
Asn	Gly	Ala	Asn	Gln	Ser	Leu	Ala	Thr	Pro	Ala	Asn	Asp	Ala	Ala	Gly
Gln	Val	Glu	Val	Ala	Phe	Val	Leu	Ala	Gln	Ala	Ser	Ala	Val	Gln	Phe
Asp	Ile	Glu	Ala	Asn	Phe	Ala	Asn	Ala	Glu	Asp	Asp	Ser	Phe	Tyr	Phe
Gln	Leu	Asn	Gly	Gly	Ala	Trp	Gln	Thr	Phe	Asn	Asn	Ala	Thr	Thr	Val
Gly	Trp	Gln	Thr	Leu	Pro	Val	Ala	Ser	Leu	Gly	Asn	Leu	Ala	Ala	Gly
Arg	His	Val	Leu	Thr	Leu	Leu	Arg	Arg	Glu	Asp	Gly	Ala	Lys	Leu	Gly
Lys	Val	Val	Leu	Ser	Ala	Ala	Gln	Ser	Ser	Ile	Ser	Arg	Ala	Thr	Pro
Val	Ala	Tyr	Ala	Ser	Pro	Asn	Asp	Val	Ala	Asn	Leu	Phe	Lys	Leu	Ala
Ser	Phe	Pro	Ile	Gly	Val	Ala	Val	Ser	Ala	Gly	Asn	Glu	Gly	Asp	Ser

595 600 605
 Leu Leu Arg Ser Gly Thr Arg Ala Ala Ala Glu Arg Ala Leu Thr Glu
 610 615 620
 Lys His Phe Asn Ser Leu Val Ala Gly Asn Ile Met Lys Met Ser Tyr
 625 630 635 640
 Leu His Pro Ala Glu Asn Thr Tyr Thr Phe Thr Gln Ala Asp Ala Leu
 645 650 655
 Ala Asp Tyr Ala Lys Ser Lys Gly Met Val Leu His Gly His Ala Leu
 660 665 670
 Val Trp His Ala Asp Tyr Gln Val Pro Asn Trp Met Lys Asn Tyr Thr
 675 680 685
 Gly Asp Trp Ser Lys Met Leu Glu Ala His Val Thr Thr Val Ala Lys
 690 695 700
 His Tyr Ala Gly Lys Val Val Ser Trp Asp Val Val Asn Glu Ala Leu
 705 710 715 720
 Ala Asp Gly Asn Ala Thr Ala Thr Lys Gly Phe Arg Ala Thr Asp Ser
 725 730 735 740
 Ile Phe Tyr Gln Lys Met Gly Ser Ser Phe Ile Glu Lys Ala Phe Ile
 745 750
 Ala Ala Arg Ala Ala Asp Pro Asn Ala Asp Leu Tyr Tyr Asn Asp Tyr
 755 760 765
 Gly Met Glu Gly Gly Asn Ser Lys Phe Asn Tyr Cys Met Ala Met Val
 770 775 780
 Asp Asp Phe Gln Lys Arg Gly Ile Pro Ile Asp Gly Ile Gly Phe Gln
 785 790 795 800
 Met His Ile Asn Ile Asp Trp Pro Ser Ser Ala Gln Ile Arg Ala Val
 805 810 815
 Phe Ser Glu Val Val Lys Arg Gly Leu Lys Val Arg Ile Ser Glu Leu
 820 825 830
 Asp Ile Pro Val Asn Thr Thr Ala Gly Arg Phe Ala Ser Leu Asn Ala
 835 840 845
 Thr Ala Asn Glu Leu Gln Lys Lys Lys Tyr Arg Glu Val Val Ala Ala
 850 855 860
 Tyr Leu Asp Val Val Pro Glu Leu Arg Gly Ile Thr Val Trp
 865 870 875 880
 Gly Leu Ser Asp Asn Gly Ser Trp Leu Val Thr Pro Thr Lys Pro Asp
 885 890 895 900
 Trp Pro Leu Leu Phe Asp Ala Asp Leu Lys Ala Lys Asp Ala Leu Ser
 900 905 910
 Gly Phe Ala Asp Ala Leu Arg Gly Val Arg
 915 920

<210> 39
 <211> 1143
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 39
 atgaaaaaaa cgattgcaca ttccacctta tggatagtgt tttttctctt cacttcctgt 60
 gctgttacgg cgcagaagaa tgcaaagaat acaagagtaa aaccactac cctaaaagag 120
 gcttaccaag gtaaatctta tatcgggtact gcatgaact tgagacagat tcacggagat 180
 gatccccaat ctgaaaatat tatcaaaaaa cagttcaatt ccatagttgc cgaaaactgc 240
 atgaagagta tgtatcttca gccggaggaa ggaaaatttt tcttcgatga tgcggacaag 300
 ttgtgtgatt ttgtgtctca gaacaatatg ttcatcattg ggcattgtct gatttggcat 360
 tcgcaggcgc caaatgggt ttccaccgat gagaatggaa acacgggttc tccagaagt 420
 cttaaacaaa ggaatgaaagc ccatattacc gccgtcgttt cccgttaca agggaaaatc 480
 aaaggttggg atgtgttgaa cgaagccatt atggaagatg gttcttaccg taaaagcaaa 540
 ttttacgaga ttttgggaga agaatttatt ccgttggcat ttcagtatgc gcatgaagca 600
 gatcctgat cagaacttta ttacaacgat tataacgaat ggtatcccgg aaaaagagct 660
 acggtgacca agataatccg cgatttcaaa actagaggaa tccgcatcga tgccatcgga 720
 atgcaggctc atttcgggat ggattcgtccc actgtagaag agtatgaaca aactattcag 780
 ggctatataa aagaaggcgt gaaagtcaat attacggaac tcgatttgag tccacttcct 840
 tctccttggg gaacttccgc caatgttgcc gatacgcagc aatatcagga aaaaatgaat 900
 ccatacaca aaggacttcc tgcagatgtt gaaaaagcat gggaaaaccg ttatgtggat 960
 tttttcaaac tgttctctaa atatcatcag catattgagc gtgttacgtt ttggggcggt 1020
 agcgatatcg atttcctgaa gaacgatatt ccggttaagag gacgtaccga ttatccacta 1080

ccgtttaacc gtcaatatca agcaaaacct ttggttcaga aattaataga ttttaacaaaa 1140
tag 1143

<210> 40
<211> 380
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(24)

<400> 40
Met Lys Lys Thr Ile Ala His Phe Thr Leu Trp Ile Val Phe Phe Leu
1 5 10 15
Phe Thr Ser Cys Ala Val Thr Ala Gln Lys Asn Ala Lys Asn Thr Arg
20 25 30
Val Lys Pro Thr Thr Leu Lys Glu Ala Tyr Gln Gly Lys Phe Tyr Ile
35 40 45
Gly Thr Ala Met Asn Leu Arg Gln Ile His Gly Asp Asp Pro Gln Ser
50 55 60
Glu Asn Ile Ile Lys Lys Gln Phe Asn Ser Ile Val Ala Glu Asn Cys
65 70 75 80
Met Lys Ser Met Tyr Leu Gln Pro Glu Glu Gly Lys Phe Phe Phe Asp
85 90 95
Asp Ala Asp Lys Phe Val Asp Phe Gly Leu Gln Asn Asn Met Phe Ile
100 105 110
Ile Gly His Cys Leu Ile Trp His Ser Gln Ala Pro Lys Trp Phe Phe
115 120 125
Thr Asp Glu Asn Gly Asn Thr Val Ser Pro Glu Val Leu Lys Gln Arg
130 135 140
Met Lys Ala His Ile Thr Ala Val Val Ser Arg Tyr Lys Gly Lys Ile
145 150 155 160
Lys Gly Trp Asp Val Val Asn Glu Ala Ile Met Glu Asp Gly Ser Tyr
165 170 175
Arg Lys Ser Lys Phe Tyr Glu Ile Leu Gly Glu Glu Phe Ile Pro Leu
180 185 190
Ala Phe Gln Tyr Ala His Glu Ala Asp Pro Asp Ala Glu Leu Tyr Tyr
195 200 205
Asn Asp Tyr Asn Glu Trp Tyr Pro Gly Lys Arg Ala Thr Val Thr Lys
210 215 220
Ile Ile Arg Asp Phe Lys Thr Arg Gly Ile Arg Ile Asp Ala Ile Gly
225 230 235 240
Met Gln Ala His Phe Gly Met Asp Ser Pro Thr Val Glu Glu Tyr Glu
245 250 255
Gln Thr Ile Gln Gly Tyr Ile Lys Glu Gly Val Lys Val Asn Ile Thr
260 265 270
Glu Leu Asp Leu Ser Pro Leu Pro Ser Pro Trp Gly Thr Ser Ala Asn
275 280 285
Val Ala Asp Thr Gln Gln Tyr Gln Glu Lys Met Asn Pro Tyr Thr Lys
290 295 300
Gly Leu Pro Ala Asp Val Glu Lys Ala Trp Glu Asn Arg Tyr Val Asp
305 310 315 320
Phe Phe Lys Leu Phe Leu Lys Tyr His Gln His Ile Glu Arg Val Thr
325 330 335
Phe Trp Gly Val Ser Asp Ile Asp Ser Trp Lys Asn Asp Phe Pro Val
340 345 350
Arg Gly Arg Thr Asp Tyr Pro Leu Pro Phe Asn Arg Gln Tyr Gln Ala
355 360 365
Lys Pro Leu Val Gln Lys Leu Ile Asp Leu Thr Lys
370 375 380

<210> 41
<211> 1893
<212> DNA
<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 41

atgatccatc	aacaaaagcc	caaccaagac	atcggtaggc	tattcaagcg	cagctgcagc	60
tttgtcggta	ttagcgcggc	actggctgtt	ttctcacaca	ccgcaagtgc	agcctgtact	120
tacaacattg	ataaccaatg	gggcagcggg	tttgtcgcta	gtattactgt	aaagaatgac	180
actggtgcaa	ccgtcaataa	ctggagtgtg	aattggcaat	atgccaaaca	tcgcatcacc	240
aatggttggg	gtgcaaattt	ctctggcagc	aatccttaca	ccgccaccaa	tatgagctgg	300
aacggtagca	ttgccgctgg	ccagtcggtg	acttttgggt	tccagggcaa	cactaacagc	360
aataccgttg	agcgcccggg	ggttaacggg	tcactgtgcy	gtactgcaac	aacctcttca	420
gttcgctcca	gcgtggctgc	gacgtcttcc	agtcgctcca	gtgttgcgcc	cagctcgatt	480
cctgcttcca	gcactccgcy	ttcaagcaca	cctgccacct	cttcttctgc	ttccagcttc	540
tcagtaccgg	ccaataattt	tgcgcagaat	ggcggcgctg	aatctgggtt	gaccaactgg	600
ggtacgactg	cgggcaccgt	gactcgctct	actgccgata	aacacagcgg	tacagccagt	660
gccttaatta	ccggccgcac	tgctgcctgg	aatgggttga	cgtttaatgt	gggcgcattg	720
accaacggca	accagtacca	agtcaacgtg	tgggtgaaat	tggctccagg	tacgcccgcg	780
agcgtactga	ccttaaccgg	taagcgtgta	gacgatagcy	atactactac	ctacaacgaa	840
tacacacgcy	tagcgactgt	gactgcctct	gccaatgagt	ggcgtttgct	ggaagggttac	900
tacacccaat	ctggcagcac	tgcatctccg	catttccatta	tcgaagcaac	ggatactact	960
gccagtttatt	acgcggatga	tttcgccatc	ggcggtcaag	tcgtacaagt	tccaagcagc	1020
agctcacgca	gctcaagcag	tgctccggcy	gctagaaaat	tcctcggcaa	catcaccacc	1080
tcgggtgcag	tgagatccga	ctttactcgt	tactggaacc	aaattacacc	agagaacgaa	1140
ggtaagtggg	gttccggttg	aggtactcgc	aaccagtaca	actgggcacc	gctggatcgt	1200
atttatgctt	acgctcgcca	aaataatatt	ccggtaaaag	ctcacacggt	tgtgtggggg	1260
gcgcaatcac	ccgcgtggct	caataactta	agcggaccgg	aagtgcgtgt	tgaaattgaa	1320
caatggattc	gcgattactg	tactcgttac	cctgacacgg	cgatgattga	cgtagtgaac	1380
gaagcggttc	ctggccatca	accggcaggt	tatgcacaac	gagcatttgg	caataactgg	1440
atccaacgcy	tgttccaatt	ggctcgccaa	tattgcccta	actcgatcct	gatcctgaat	1500
gattacaaca	atatccggtg	gcagcacaat	gagtttattg	cccttgcaaa	agctcaaggc	1560
aattatattg	atgcagtcgg	cctgcaggcy	catgaactga	agggatatgac	agcggcgcaa	1620
gtcaaaaccg	caatcgacaa	tatttggaac	caagtgggca	agcccatcta	catttctgaa	1680
tacgacattg	gcgataacaa	tgaccaggtt	caattgcaga	atttccaggc	gcatttccct	1740
gtattctggg	accatccgca	tgttaaaggc	atcaccattt	ggggttatgt	caatggcaga	1800
acttggattg	aaggctcggg	cctgatttct	gacaacggaa	caccgcgccc	cgcaatgact	1860
tggttgctga	ataactatat	caataagcag	taa			1893

<210> 42

<211> 630

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(37)

<400> 42

Met	Ile	His	Gln	Gln	Lys	Pro	Asn	Gln	Asp	Ile	Gly	Arg	Leu	Phe	Lys
1			5					10						15	
Arg	Ser	Cys	Ser	Phe	Val	Gly	Ile	Ser	Ala	Ala	Leu	Ala	Val	Phe	Ser
			20					25					30		
His	Thr	Ala	Ser	Ala	Ala	Cys	Thr	Tyr	Asn	Ile	Asp	Asn	Gln	Trp	Gly
		35					40					45			
Ser	Gly	Phe	Val	Ala	Ser	Ile	Thr	Val	Lys	Asn	Asp	Thr	Gly	Ala	Thr
	50					55					60				
Val	Asn	Asn	Trp	Ser	Val	Asn	Trp	Gln	Tyr	Ala	Asn	Asn	Arg	Ile	Thr
65					70				75					80	
Asn	Gly	Trp	Ser	Ala	Asn	Phe	Ser	Gly	Ser	Asn	Pro	Tyr	Thr	Ala	Thr
			85					90						95	
Asn	Met	Ser	Trp	Asn	Gly	Ser	Ile	Ala	Ala	Gly	Gln	Ser	Val	Thr	Phe
		100						105					110		
Gly	Phe	Gln	Gly	Asn	Thr	Asn	Ser	Asn	Thr	Val	Glu	Arg	Pro	Val	Val
		115				120					125				
Asn	Gly	Ser	Leu	Cys	Gly	Thr	Ala	Thr	Thr	Ser	Ser	Val	Arg	Ser	Ser
	130					135					140				
Val	Ala	Ala	Thr	Ser	Ser	Ser	Arg	Ser	Ser	Val	Ala	Pro	Ser	Ser	Ile
145					150				155						160

Pro Ala Ser Ser Thr Pro Arg Ser Ser Thr Pro Ala Thr Ser Ser Ser
 165 170 175
 Ala Ser Ser Phe Ser Val Pro Ala Asn Asn Phe Ala Gln Asn Gly Gly
 180 185 190
 Val Glu Ser Gly Leu Thr Asn Trp Gly Thr Thr Ala Gly Thr Val Thr
 195 200 205
 Arg Ser Thr Ala Asp Lys His Ser Gly Thr Ala Ser Ala Leu Ile Thr
 210 215 220
 Gly Arg Thr Ala Ala Trp Asn Gly Leu Thr Phe Asn Val Gly Ala Leu
 225 230 235 240
 Thr Asn Gly Asn Gln Tyr Gln Val Asn Val Trp Val Lys Leu Ala Pro
 245 250 255
 Gly Thr Pro Asp Ser Val Leu Thr Leu Thr Gly Lys Arg Val Asp Asp
 260 265 270
 Ser Asp Thr Thr Thr Tyr Asn Glu Tyr Thr Arg Val Ala Thr Val Thr
 275 280 285
 Ala Ser Ala Asn Glu Trp Arg Leu Leu Glu Gly Tyr Tyr Thr Gln Ser
 290 295 300
 Gly Ser Thr Ala Phe Gln His Phe Ile Ile Glu Ala Thr Asp Thr Thr
 305 310 315 320
 Ala Ser Tyr Tyr Ala Asp Asp Phe Ala Ile Gly Gly Gln Val Val Gln
 325 330 335
 Val Pro Ser Ser Ser Ser Arg Ser Ser Ser Ser Ala Pro Ala Ala Arg
 340 345 350
 Lys Phe Ile Gly Asn Ile Thr Thr Ser Gly Ala Val Arg Ser Asp Phe
 355 360 365
 Thr Arg Tyr Trp Asn Gln Ile Thr Pro Glu Asn Glu Gly Lys Trp Gly
 370 375 380
 Ser Val Glu Gly Thr Arg Asn Gln Tyr Asn Trp Ala Pro Leu Asp Arg
 385 390 395 400
 Ile Tyr Ala Tyr Ala Arg Gln Asn Asn Ile Pro Val Lys Ala His Thr
 405 410 415
 Phe Val Trp Gly Ala Gln Ser Pro Ala Trp Leu Asn Asn Leu Ser Gly
 420 425 430
 Pro Glu Val Ala Val Glu Ile Glu Gln Trp Ile Arg Asp Tyr Cys Thr
 435 440 445
 Arg Tyr Pro Asp Thr Ala Met Ile Asp Val Val Asn Glu Ala Val Pro
 450 455 460
 Gly His Gln Pro Ala Gly Tyr Ala Gln Arg Ala Phe Gly Asn Asn Trp
 465 470 475 480
 Ile Gln Arg Val Phe Gln Leu Ala Arg Gln Tyr Cys Pro Asn Ser Ile
 485 490 495
 Leu Ile Leu Asn Asp Tyr Asn Asn Ile Arg Trp Gln His Asn Glu Phe
 500 505 510
 Ile Ala Leu Ala Lys Ala Gln Gly Asn Tyr Ile Asp Ala Val Gly Leu
 515 520 525
 Gln Ala His Glu Leu Lys Gly Met Thr Ala Ala Gln Val Lys Thr Ala
 530 535 540
 Ile Asp Asn Ile Trp Asn Gln Val Gly Lys Pro Ile Tyr Ile Ser Glu
 545 550 555 560
 Tyr Asp Ile Gly Asp Asn Asn Asp Gln Val Gln Leu Gln Asn Phe Gln
 565 570 575
 Ala His Phe Pro Val Phe Trp Asp His Pro His Val Lys Gly Ile Thr
 580 585 590
 Ile Trp Gly Tyr Val Asn Gly Arg Thr Trp Ile Glu Gly Ser Gly Leu
 595 600 605
 Ile Ser Asp Asn Gly Thr Pro Arg Pro Ala Met Thr Trp Leu Leu Asn
 610 615 620
 Asn Tyr Ile Asn Lys Gln 630

<210> 43
 <211> 1011
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

```

<400> 43
atgcaaaca atattaaagg aaataacatt ccatcattac acgaagttaa tcaagatcac      60
tttttgatag gtgcagcagt taatccaaaa acattagact cacagcagga tttattgaga      120
aaacacttta acagtattac agctgaaaat gaaatgaaat ttgaagaatt gcaaccagaa      180
cctggccatt tcacgttttg tgtagcagat gaaatcgatt catttgcaaa agaaaatgga      240
atgaaagtta gaggacatac attagtttgg cataatcaaa cgcttgattg gatgtttttg      300
aatgaagatg gatctgtcac agatcgagaa acgcttctag aaagaatgaa attacacatt      360
acaacagtta tgcagcatta caaagggtcaa gcttatttgc gggatgttgt aaatgagggtg      420
attgctgacg aggggtacaga gttattccgt aaatctaaat ggactgaaat tattgggtgat      480
gattttgtag aaaaggcatt tgaatatgca catgaggctg atccagaagc tttactattc      540
tacaatgact ataatgaatc ccatcccaat aagcgtgaga aaattttcac acttgtaaaa      600
ggattagtgt ataaggggat acctattcat ggaatcggtt tacaagcaca ttggaattta      660
acaggacctt cttatgaaga tattagagca gcactcgaga aatatgctac attgggattg      720
gaaatacacc ttaccgaatt ggatgtttct gtttttaatt atgaagatcg aagaacagat      780
ttaacagaac caactaaaga tatgcaagcg cttcaagcgg agcgttatac agaattattc      840
aagatatgga gagaatatag tcatgtaatc agttcgatta ctttttgggg agctgcagat      900
gattatactt ggtttagatga ttttcctgtc aaaggaagaa aaaactggcc atttgttttt      960
gatgaaaacc aagagccaaa agagtcattt tggaatatta ttgactttta a      1011

```

```

<210> 44
<211> 336
<212> PRT
<213> Unknown

```

```

<220>
<223> obtained from an environmental sample

```

```

<400> 44
Met Gln Thr Asn Ile Lys Gly Asn Asn Ile Pro Ser Leu His Glu Val
1      5      10      15
Tyr Gln Asp His Phe Leu Ile Gly Ala Ala Val Asn Pro Lys Thr Leu
20      25      30
Asp Ser Gln Gln Asp Leu Leu Arg Lys His Phe Asn Ser Ile Thr Ala
35      40      45
Glu Asn Glu Met Lys Phe Glu Glu Leu Gln Pro Glu Pro Gly His Phe
50      55      60
Thr Phe Gly Val Ala Asp Glu Ile Val Ser Phe Ala Lys Glu Asn Gly
65      70      75      80
Met Lys Val Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro Asp
85      90      95
Trp Met Phe Leu Asn Glu Asp Gly Ser Val Thr Asp Arg Glu Thr Leu
100      105      110
Leu Glu Arg Met Lys Leu His Ile Thr Thr Val Met Gln His Tyr Lys
115      120      125
Gly Gln Ala Tyr Cys Trp Asp Val Val Asn Glu Val Ile Ala Asp Glu
130      135      140
Gly Thr Glu Leu Phe Arg Lys Ser Lys Trp Thr Glu Ile Ile Gly Asp
145      150      155      160
Asp Phe Val Glu Lys Ala Phe Glu Tyr Ala His Glu Ala Asp Pro Glu
165      170      175
Ala Leu Leu Phe Tyr Asn Asp Tyr Asn Glu Ser His Pro Asn Lys Arg
180      185      190
Glu Lys Ile Phe Thr Leu Val Lys Gly Leu Val Asp Lys Gly Ile Pro
195      200      205
Ile His Gly Ile Gly Leu Gln Ala His Trp Asn Leu Thr Gly Pro Ser
210      215      220
Tyr Glu Asp Ile Arg Ala Ala Leu Glu Lys Tyr Ala Thr Leu Gly Leu
225      230      235      240
Glu Ile His Leu Thr Glu Leu Asp Val Ser Val Phe Asn Tyr Glu Asp
245      250      255
Arg Arg Thr Asp Leu Thr Glu Pro Thr Lys Asp Met Gln Ala Leu Gln
260      265      270
Ala Glu Arg Tyr Thr Glu Leu Phe Lys Ile Leu Arg Glu Tyr Ser His
275      280      285
Val Ile Ser Ser Ile Thr Phe Trp Gly Ala Ala Asp Tyr Thr Trp
290      295      300
Leu Asp Asp Phe Pro Val Lys Gly Arg Lys Asn Trp Pro Phe Val Phe
305      310      315      320
Asp Glu Asn Gln Glu Pro Lys Glu Ser Phe Trp Asn Ile Ile Asp Phe

```

335

<220>
<223> obtained from an environmental sample

<210>	46
<211>	378
<212>	PRT
<213>	Unknown

<220>
<223> obtained from an environmental sample

<221> SIGNAL
<222> (1)...(25)

Met 1	Lys 46	Ile	Ser	Arg 5	Arg	Gln	Leu	Leu	Ala 10	Met	Gly	Gly	Ala 15	Ala	Ala
Thr	Leu	Ala	Ser 20	Ala	Lys	Leu	Phe	Ala 25	Ala	Glu	Lys	Ala 30	Ala	Ala	Ala
Thr	Gly	Leu 35	Lys	Asp	Ala	Tyr	Lys 40	Asn	Asp	Phe	Leu	Ile 45	Gly	Ala	Ala
Leu	Asn 50	Thr	Gln	Ile	Val 55	Asp	Gly	Lys	Asp	Pro	Lys 60	Leu	Thr	Ala	Leu
Ile 65	Thr	Lys	Glu	Phe	Asn 70	Ser	Ile	Thr	Ala 75	Glu	Asn	Cys	Gln	Lys	Trp 80
Glu	Arg	Leu	Arg	Asn 85	Glu	Lys	Asp	Gly	Ser 90	Trp	Glu	Trp	Lys	Asp 95	Ser
Asp	Ala	Phe	Val 100	Asn	Phe	Gly	Val	Ala 105	His	Asn	Met	His	Ile 110	Val	Gly
His	Thr	Leu 115	Gly	Trp	His	Ser	Gln 120	Ile	Pro	Asp	Ser	Val 125	Phe	Lys	Asn
Lys	Asp 130	Gly	Ser	Tyr	Ile	Ser 135	Lys	Glu	Ala	Leu	Ala 140	Lys	Lys	Gln	Gln
Glu 145	His	Ile	Thr	Thr	Leu 150	Val	Asp	Arg	Tyr	Lys 155	Gly	Lys	Ile	Ala	Ala 160
Trp	Asp	Val	Val	Asn 165	Glu	Ala	Met	Gly	Asp 170	Asp	Asn	Lys	Met	Arg 175	Ala
Ser	His	Trp	Tyr 180	Asn	Ile	Met	Gly	Asp 185	Asp	Phe	Leu	Val	Asn 190	Ala	Phe
Lys	Leu	Ala 195	His	Glu	Thr	Asp	Pro 200	Lys	Ala	His	Leu	Met 205	Tyr	Asn	Asp

Tyr Asn Asn Glu Arg Pro Glu Lys Arg Ala Ala Thr Val Asp Met Leu
 210 215 220
 Lys Arg Leu Leu Lys Leu Gly Ala Pro Ile His Gly Leu Gly Met Gln
 225 230 235 240
 Ala His Ile Gly Leu Asp Ala Asp Met Lys Asn Phe Glu Asp Ser Ile
 245 250 255
 Val Ala Tyr Ser Glu Leu Gly Leu Arg Ile His Leu Thr Glu Leu Asp
 260 265 270
 Ile Asp Val Leu Pro Ser Val Trp Asn Leu Pro Val Ala Glu Val Ser
 275 280 285
 Thr Arg Phe Glu Tyr Lys Pro Glu Arg Asp Pro Tyr Ile Lys Gly Leu
 290 295 300
 Pro Lys Glu Ile Asp Glu Lys Leu Ala Lys Ala Tyr Glu Ser Leu Phe
 305 310 315 320
 Lys Ile Leu Leu Lys His Lys Asp Lys Val Asp Arg Val Thr Phe Trp
 325 330 335
 Gly Val Ser Asp Asp Ala Ser Trp Leu Asn Gly Phe Pro Ile Pro Gly
 340 345 350
 Arg Thr Asn Tyr Pro Leu Leu Phe Asp Arg Lys Gln Gln Pro Lys Ala
 355 360 365
 Ala Tyr Phe Arg Leu Leu Asp Leu Lys Arg
 370 375

<210> 47
 <211> 1137
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 47
 atgaaaagaa taaagattct gaattcgatt gtattagctt taatcctggc gatcatcctg 60
 ccgggatgtt ccaatgcaca gaagagcgag ccggtgctga aagatgccct ttcgggaaaa 120
 ttftacatcg gggctgctct caataccccc caaattacgg gccgggatac cttgtccatg 180
 aaaaagggtca ccagacattt taactccatc gtagctgaga actgcatgaa aagcggggag 240
 atccagcgga ccgaagggga gtttgatttc agtcttgccg accagtttgt cgcgttcggc 300
 gaaaaacaca acatgcacat tgtggggcat accctgatat ggcattcaca ggcgccgcgc 360
 tggtttttca ccggtgcaga cggaaacgaa gtcagccggg aggtactgat tgagcgcgatg 420
 aagaaccata ttatatacgg cgtggggcgt tacaaggcc gtgtccacgg ctgggatgtg 480
 gtcaacgaag ccattgaaga caacggctca tggcgcaaca gcaagtttta ccagatctta 540
 ggtgacgagt ttgtggaact ggcctttaaa ttgcccag aagccgaccc ggatgccgaa 600
 cttactata acgactactc catggcatta gaaggcagga gaaatggcgt tatcagaatg 660
 gtgaagaacc ttcatgtcaa gggactcaaa attgacggt tccgcatgca ggggcatctg 720
 ctcatggact cgcccacgct ggaagcttat gaagaaagta tcctggccta ttccggactg 780
 ggcgttaagg tgatgatcac ggaactcgat ttgtctgcgc tgccatggcc agcccgtcag 840
 cagggaagccg atattgccct gagggctgag tatgaggcac ggatgaatcc ttacaccgaa 900
 ggtttaaccg attcatcttc cgtggcatgg aatcagcgga tgggcgattt cttctctctt 960
 ttctgaagc accaggacaa aatcagcagg gttacccttt ggggggtcac cgataaccaa 1020
 tcctggaaaa ataactttcc gatgagagga aggacagact acccgttgct ttttgaccgg 1080
 aattaccaac ccaaaccggt ggtggaaaga atcatcaaa aagcgaaagc aaaataa 1137

<210> 48
 <211> 378
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(26)

<400> 48
 Met Lys Arg Ile Lys Ile Leu Asn Ser Ile Val Leu Ala Leu Ile Leu
 1 5 10 15
 Ala Ile Ile Leu Pro Gly Cys Ser Asn Ala Gln Lys Ser Glu Pro Val
 20 25 30
 Leu Lys Asp Ala Leu Ser Gly Lys Phe Tyr Ile Gly Ala Ala Leu Asn
 Page 37

35 40 45
 Thr Pro Gln Ile Thr Gly Arg Asp Thr Leu Ser Met Lys Met Val Thr
 50 55 60
 Arg His Phe Asn Ser Ile Val Ala Glu Asn Cys Met Lys Ser Gly Glu
 65 70 75 80
 Ile Gln Arg Thr Glu Gly Glu Phe Asp Phe Ser Leu Ala Asp Gln Phe
 85 90 95
 Val Ala Phe Gly Glu Lys His Asn Met His Ile Val Gly His Thr Leu
 100 105 110
 Ile Trp His Ser Gln Ala Pro Arg Trp Phe Phe Thr Gly Ala Asp Gly
 115 120 125
 Asn Glu Val Ser Arg Glu Val Leu Ile Glu Arg Met Lys Asn His Ile
 130 135 140
 Tyr Thr Val Val Gly Arg Tyr Lys Gly Arg Val His Gly Trp Asp Val
 145 150 155 160
 Val Asn Glu Ala Ile Glu Asp Asn Gly Ser Trp Arg Asn Ser Lys Phe
 165 170 175
 Tyr Gln Ile Leu Gly Asp Glu Phe Val Glu Leu Ala Phe Lys Phe Ala
 180 185 190
 Ala Glu Ala Asp Pro Asp Ala Glu Leu Tyr Tyr Asn Asp Tyr Ser Met
 195 200 205
 Ala Leu Glu Gly Arg Arg Asn Gly Val Ile Arg Met Val Lys Asn Leu
 210 215 220
 Gln Ser Lys Gly Leu Lys Ile Asp Gly Ile Gly Met Gln Gly His Leu
 225 230 235 240
 Leu Met Asp Ser Pro Thr Leu Glu Ala Tyr Glu Glu Ser Ile Leu Ala
 245 250 255
 Tyr Ser Gly Leu Gly Val Lys Val Met Ile Thr Glu Leu Asp Leu Ser
 260 265 270
 Ala Leu Pro Trp Pro Ala Arg Gln Gln Gly Ala Asp Ile Ala Leu Arg
 275 280 285
 Ala Glu Tyr Glu Ala Arg Met Asn Pro Tyr Thr Glu Gly Leu Thr Asp
 290 295 300
 Ser Ala Ser Val Ala Trp Asn Gln Arg Met Gly Asp Phe Phe Ser Leu
 305 310 315 320
 Phe Leu Lys His Gln Asp Lys Ile Ser Arg Val Thr Leu Trp Gly Val
 325 330 335
 Thr Asp Asn Gln Ser Trp Lys Asn Asn Phe Pro Met Arg Gly Arg Thr
 340 345 350
 Asp Tyr Pro Leu Leu Phe Asp Arg Asn Tyr Gln Pro Lys Pro Val Val
 355 360 365
 Glu Arg Ile Ile Lys Glu Ala Lys Ala Lys
 370 375

<210> 49
 <211> 996
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 49
 atgaacagct ccctccctc cctccgcat gtattcgcga atgatttccg catcggggcg 60
 gcggtcaatc ctgtgacgat cgagatgcaa aaacagttgt tgatcgatca tgtcaacagt 120
 attacggcag agaaccatat gaagtttgag catcttcagc cggaagaagg gaaatttacc 180
 tttcaggaag cggatcggat tgtggatttt gcttgttcgc accgaatggc gggttcgaggg 240
 cacacacttg tatggcacia ccagactccg gattgggtgt ttcaagatgg tcaaggccat 300
 ttcgtcagtc gggatgtgtt gcttgagcgg atgaaatgtc acatttcaac tgttgtacgg 360
 cgatacaagg gaaaaatata ttgttgggat gtcatacaac aagcggtagc cgacgaagga 420
 gacgaattgt tgaggccgtc gaagtggcga caaatcatcg gggacgattt tatggaacaa 480
 gcattttctt acgcttatga agctgaccca gatgcactgc ttttttacia tgactataat 540
 gaatgttttc cggaaaagag agaaaaaatt ttgacacttg tcaaactcgt gcgtgataaa 600
 ggcattccga ttcatggcat cggcatgcag gcgcactgga gcctgacctg cccgtcgctt 660
 gatgaaattc gtgcggcgat tgaacggtat gcgtcccttg gtgttgttct tcatattacg 720
 gaactcgatg tatccatgtt tgaatttcac gatcgtcgaa ccgatttggc tgtcccagc 780
 aacgaaatga tcgaacagca agcagaacgg tatgggcaaa tttttgctt gttaaaggag 840
 tatcgcgatg ttattcaaa tgctcacattt tgggggaatt ctgatgacca tacatggctc 900
 gataactttc cagtgcacgg gagaaaaaac tggccgcttt tgttcgatga acagcataaa 960

ccgaaaccag ctttttggcg ggcagtgagt gtctga

996

<210> 50
 <211> 331
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 50
 Met Asn Ser Ser Leu Pro Ser Leu Arg Asp Val Phe Ala Asn Asp Phe
 1 5 10 15
 Arg Ile Gly Ala Ala Val Asn Pro Val Thr Ile Glu Met Gln Lys Gln
 20 25 30
 Leu Leu Ile Asp His Val Asn Ser Ile Thr Ala Glu Asn His Met Lys
 35 40 45
 Phe Glu His Leu Gln Pro Glu Glu Gly Lys Phe Thr Phe Gln Glu Ala
 50 55 60
 Asp Arg Ile Val Asp Phe Ala Cys Ser His Arg Met Ala Val Arg Gly
 65 70 75 80
 His Thr Leu Val Trp His Asn Gln Thr Pro Asp Trp Val Phe Gln Asp
 85 90 95
 Gly Gln Gly His Phe Val Ser Arg Asp Val Leu Leu Glu Arg Met Lys
 100 105 110
 Cys His Ile Ser Thr Val Val Arg Arg Tyr Lys Gly Lys Ile Tyr Cys
 115 120 125
 Trp Asp Val Ile Asn Glu Ala Val Ala Asp Glu Gly Asp Glu Leu Leu
 130 135 140
 Arg Pro Ser Lys Trp Arg Gln Ile Ile Gly Asp Asp Phe Met Glu Gln
 145 150 155 160
 Ala Phe Leu Tyr Ala Tyr Glu Ala Asp Pro Asp Ala Leu Leu Phe Tyr
 165 170 175
 Asn Asp Tyr Asn Glu Cys Phe Pro Glu Lys Arg Glu Lys Ile Phe Ala
 180 185 190
 Leu Val Lys Ser Leu Arg Asp Lys Gly Ile Pro Ile His Gly Ile Gly
 195 200 205
 Met Gln Ala His Trp Ser Leu Thr Arg Pro Ser Leu Asp Glu Ile Arg
 210 215 220
 Ala Ala Ile Glu Arg Tyr Ala Ser Leu Gly Val Val Leu His Ile Thr
 225 230 235 240
 Glu Leu Asp Val Ser Met Phe Glu Phe His Asp Arg Arg Thr Asp Leu
 245 250 255
 Ala Val Pro Thr Asn Glu Met Ile Glu Gln Gln Ala Glu Arg Tyr Gly
 260 265 270
 Gln Ile Phe Ala Leu Phe Lys Glu Tyr Arg Asp Val Ile Gln Ser Val
 275 280 285
 Thr Phe Trp Gly Ile Ala Asp Asp His Thr Trp Leu Asp Asn Phe Pro
 290 295 300
 Val His Gly Arg Lys Asn Trp Pro Leu Leu Phe Asp Glu Gln His Lys
 305 310 315 320
 Pro Lys Pro Ala Phe Trp Arg Ala Val Ser Val
 325 330

<210> 51
 <211> 3162
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 51
 atgagagggg aaagcaaaaa gggatttctg aacatctcag aagctgtact tgttgggaatt 60
 ttagcagggt ttcttggagt tcttctcgca gctacggggg ttttgagttt tgggtgaaca 120
 gcgtcttcgt ctcttgaaac ggtgttcacc ttgagtttcg agggaaacaac gcaaggtgtc 180
 aatccctttg gaaaagaagt agttctcaca gcttctcaag atgtagcagc cgatggcgaa 240
 tattcattga aagtagagaa tagaacttcc ggctgggatg gagttgagat cgatttaacg 300
 gaaaaagtag aagcgaacaa agattatctg ttgtctttct acgtctatca aacatctgac 360

tcaccccaac	tttttgaagt	ccttgcaaga	acagaagacg	ggaaagggtga	aaaatacgaa	420
acccttaccg	acaagggtgt	agtatcgaac	tactggaaag	aaatttctgt	gcccttttcc	480
ccgagtttcg	agagtacccc	aacaaaatgt	tctttgatcg	ttgtttcacc	aaagaaccca	540
tcattcactt	tctatatattga	caaggttcaa	atttctcaaac	cgaagaagca	aggtccacaa	600
gtcattttacg	aaacatccctt	tgagagtggg	acgggaagct	ggcaagccag	aggggtctgat	660
gtgaaaaatca	aagtgcacatc	gaaagtgtgt	cattctggaa	aaaggctctt	ctatgtctcc	720
aacagacaaa	aaggctggca	tggtgtacaa	cttgacgtga	agagactctt	gagacccggg	780
aaaacgtatg	cttttgaagg	atgggtttat	caagactctg	gacaggatca	aacaattatt	840
ctgacgatgc	agagaagata	ttcttctgat	tctagcacac	aatatgagtg	gatcaaggcg	900
gtaactgttc	catcaggaca	atggacgcag	atctctggaa	cttacacaat	ccaaccaaga	960
gtgaagctgg	aggaactcat	tgtttacttt	gaagccaagg	atcccactct	tgcttcttat	1020
gtggacgatt	tcaaaataac	ggataccaca	actactgaca	tcaagctcga	gctgaagcct	1080
gaagaagaaa	ttccagctct	taaagaagtg	cttggagatt	acttcaaagt	aggtgttgcc	1140
ttacctttca	aagtttttgc	caaaccagag	gatattgtct	tcattactaa	acattttcaac	1200
agcatcactg	ccgaaaacga	aatgaaacct	gagagtctct	tggctggcgt	agaaaatgga	1260
aagttgaagt	tcagggttga	gacagcagac	aaatcagtag	aatttgacac	gcaaaacggg	1320
atgggttga	gaggtcacac	tctgggtgtg	cacaatcaaa	caccggactg	gttcttcaag	1380
gacgagaacg	gaaatctgct	ctccaaagaa	gcaatgactg	aaaggcttag	ggaatacatc	1440
cacacagtcg	tcggacactt	caaaggcaaa	gtttacgcgt	gggacgtcgt	taatgaggca	1500
gtagatccat	cccaaccaga	tggaacttaga	agatctatat	ggtacgaaat	catgggacct	1560
gactatatag	aacttgcatt	caagtttgca	agagaagcgg	accccaatgc	aaagctcttc	1620
tacaacgact	acaacaccta	ccaggagaag	aagagagaca	tcatttaca	cctcgtcaaa	1680
tccctcaaag	agaagggact	cattgacggg	atcgggtatgc	agtgtcatat	cgggtgttggg	1740
accagtgtca	aagagattga	agaggcaatc	aaaaaattca	gcaccattcc	aggtatcgaa	1800
attcatatca	cggaaactga	tataagtgtg	tacgaggatg	cgacttccaa	ttatccaaca	1860
cctccaaggg	aggctctcat	taaacaagca	cacgtaatga	gagaactctt	tgccatcttc	1920
aaaaagtaca	gcaacgtcat	aacaaacggt	actttctggg	gattgaaaga	tgattattcc	1980
tggaagaatg	cccgcagaaa	cgactggccg	ctactttttg	ataaagacta	ccaagccaaa	2040
cttgcctact	gggccatagt	cagtcctgag	gctctaccgg	tgcttccaaa	gaaatggtct	2100
atcgctacag	gtagtgtctt	ggtagttgga	atgatggatg	actcctattt	ggcttcttca	2160
cctatcaaaa	ttctcgtcga	tggccaagaa	aaactcacag	ccagagtcac	ctgggaagaa	2220
aacaaactct	tcgtctacgc	agaggtctat	gacaggacaa	gagacaaagg	aaaggacggg	2280
atcaccatct	ttgtggatcc	taaaaacttc	aaggcacctt	acttgcatga	agatgctttc	2340
tacgttacca	taaaaaccga	ctggagtgtt	gagaagagtc	gtgatgacat	agaagtccag	2400
agattcgtag	gtccaagtgg	agtaagggtac	aacgttgaat	gtgaaataac	acttcctgaa	2460
aaactccagg	aaggacagca	aatcggattt	gatatcgccg	tccaggatgg	cgataagggtc	2520
tacagctggg	ctgatatact	caatcagcag	aagctcgcaa	ccatgaacta	cggaactctc	2580
actctgcagg	gtgcggtaat	ggccacagct	aagtatggca	cacctgtgat	cgatggtgaa	2640
atagatgaca	tctgggtacac	cactgaagaa	atctcaaccg	atgttggtgt	catgggttca	2700
ctcaagaacg	caagggcaaa	agtgcagagt	ctctgggatg	aagagcacct	ctatgtgctt	2760
gccatcgtaa	ccgatcctgt	gctcaataag	gacaacacca	atccatggga	acaagactct	2820
gtagaatctt	tcatagacga	aaacaacgcc	aaaacaccgt	actatcagga	cgatgatgct	2880
caatatcggt	tcaactacct	caacgaacaa	tccttcggta	caggtgcaag	cagcaagaac	2940
ttcaagacag	ccgtgaaact	catcgatggg	ggttatcttg	ttgaggcagc	ggttaaatgg	3000
aagaccatca	aaccttcacc	aaacacagt	ataggctttg	atttccaggt	gaacgatgca	3060
aatgctcaag	gtaagagagt	tggataactt	aagtgtgtcg	atccaacgga	caacagctgg	3120
cagaatacct	ccaagtttgg	taatctcagg	ttgataaaat	ag		3162

<210> 52

<211> 1053

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(30)

<400> 52

Met	Arg	Gly	Lys	Ser	Lys	Lys	Gly	Phe	Leu	Asn	Ile	Ser	Glu	Ala	Val
1				5					10					15	
Leu	Val	Gly	Ile	Leu	Ala	Gly	Phe	Leu	Gly	Val	Leu	Leu	Ala	Ala	Thr
			20					25					30		
Gly	Val	Leu	Ser	Phe	Gly	Gly	Thr	Ala	Ser	Ser	Ser	Leu	Glu	Thr	Val
		35					40					45			
Phe	Thr	Leu	Ser	Phe	Glu	Gly	Thr	Thr	Gln	Gly	Val	Asn	Pro	Phe	Gly
	50					55					60				
Lys	Glu	Val	Val	Leu	Thr	Ala	Ser	Gln	Asp	Val	Ala	Ala	Asp	Gly	Glu

65	Tyr	Ser	Leu	Lys	Val	70	Glu	Asn	Arg	Thr	75	Ser	Gly	Trp	Asp	Gly	80	Val	Glu
	Ile	Asp	Leu	Thr	85	Glu	Lys	Val	Glu	Ala	90	Asn	Lys	Asp	Tyr	Leu	95	Leu	Ser
	Phe	Tyr	Val	Tyr	100	Gln	Thr	Ser	Asp	105	Ser	Pro	Gln	Leu	Phe	110	Glu	Val	Leu
	Ala	Arg	Thr	Glu	115	Asp	Gly	Lys	Gly	120	Glu	Lys	Tyr	Glu	Thr	125	Leu	Thr	Asp
	Lys	Val	Val	Val	130	Ser	Asn	Tyr	Trp	135	Lys	Glu	Ile	Leu	Val	140	Pro	Phe	Ser
	Pro	Ser	Phe	Glu	145	Ser	Thr	Pro	Thr	150	Lys	Cys	Ser	Leu	Ile	155	Val	Val	Ser
	Pro	Lys	Asn	Pro	165	Ser	Phe	Thr	Phe	170	Tyr	Ile	Asp	Lys	Val	175	Gln	Ile	Leu
	Lys	Pro	Lys	Lys	180	Gln	Gly	Pro	Gln	185	Val	Ile	Tyr	Glu	Thr	190	Ser	Phe	Glu
	Ser	Gly	Thr	Gly	195	Ser	Trp	Gln	Ala	200	Arg	Gly	Ser	Asp	Val	205	Lys	Ile	Lys
	Val	Thr	Ser	Lys	210	Val	Ala	His	Ser	215	Gly	Lys	Arg	Ser	Leu	220	Tyr	Val	Ser
	Asn	Arg	Gln	Lys	225	Gly	Trp	His	Gly	230	Val	Gln	Leu	Asp	Val	235	Lys	Arg	Leu
	Leu	Arg	Pro	Gly	245	Lys	Thr	Tyr	Ala	250	Phe	Glu	Gly	Trp	Val	255	Tyr	Gln	Asp
	Ser	Gly	Gln	Asp	260	Gln	Thr	Ile	Ile	265	Leu	Thr	Met	Gln	Arg	270	Arg	Tyr	Ser
	Ser	Asp	Ser	Ser	275	Thr	Gln	Tyr	Glu	280	Trp	Ile	Lys	Ala	Val	285	Thr	Val	Pro
	Ser	Gly	Gln	Trp	290	Thr	Gln	Ile	Ser	295	Gly	Thr	Tyr	Thr	Ile	300	Gln	Pro	Arg
	Val	Ser	Val	Glu	305	Glu	Leu	Ile	Val	310	Tyr	Phe	Glu	Ala	Lys	315	Asp	Pro	Thr
	Leu	Ala	Phe	Tyr	325	Val	Asp	Asp	Phe	330	Lys	Ile	Thr	Asp	Thr	335	Thr	Thr	Thr
	Asp	Ile	Lys	Leu	340	Glu	Leu	Lys	Pro	345	Glu	Glu	Glu	Ile	Pro	350	Ala	Leu	Lys
	Glu	Val	Leu	Gly	355	Asp	Tyr	Phe	Lys	360	Val	Gly	Val	Ala	Leu	365	Pro	Phe	Lys
	Val	Phe	Ala	Lys	370	Pro	Glu	Asp	Ile	375	Ala	Leu	Ile	Thr	Lys	380	His	Phe	Asn
	Ser	Ile	Thr	Ala	385	Asn	Glu	Met	Lys	390	Pro	Glu	Ser	Leu	Leu	395	Ala	Gly	Gly
	Val	Glu	Asn	Gly	405	Lys	Leu	Lys	Phe	410	Arg	Phe	Glu	Thr	Ala	415	Asp	Lys	Tyr
	Val	Glu	Phe	Ala	420	Gln	Gln	Asn	Gly	425	Met	Val	Val	Arg	Gly	430	His	Thr	Leu
	Val	Trp	His	Asn	435	Gln	Thr	Pro	Asp	440	Trp	Phe	Phe	Lys	Asp	445	Glu	Asn	Gly
	Asn	Leu	Leu	Ser	450	Glu	Ala	Met	Thr	455	Glu	Arg	Leu	Arg	Glu	460	Tyr	Ile	Ile
	His	Thr	Val	Val	465	Gly	His	Phe	Lys	470	Lys	Val	Tyr	Ala	Trp	475	Asp	Val	Val
	Val	Asn	Glu	Ala	485	Val	Asp	Pro	Ser	490	Gln	Pro	Asp	Gly	Leu	495	Arg	Arg	Ser
	Ile	Trp	Tyr	Glu	500	Ile	Met	Gly	Pro	505	Asp	Tyr	Ile	Glu	Leu	510	Ala	Phe	Lys
	Phe	Ala	Arg	Glu	515	Ala	Asp	Pro	Asn	520	Ala	Lys	Leu	Phe	Tyr	525	Asn	Asp	Tyr
	Asn	Thr	Tyr	Gln	530	Glu	Lys	Lys	Arg	535	Asp	Ile	Ile	Tyr	Asn	540	Leu	Val	Lys
	Ser	Leu	Lys	Glu	545	Lys	Gly	Leu	Ile	550	Asp	Gly	Ile	Gly	Met	555	Gln	Cys	His
	Ile	Gly	Val	Gly	565	Thr	Ser	Val	Lys	570	Glu	Ile	Glu	Glu	Ala	575	Ile	Lys	Lys
	Phe	Ser	Thr	Ile	580	Pro	Gly	Ile	Glu	585	Ile	His	Ile	Thr	Glu	590	Leu	Asp	Ile
	Ser	Val	Tyr	Glu	595	Asp	Ala	Thr	Ser	600	Asn	Tyr	Pro	Thr	Pro	605	Pro	Arg	Glu
					610					615						620			

Ala Leu Ile Lys Gln Ala His Val Met Arg Glu Leu Phe Ala Ile Phe
 625 630 635 640
 Lys Lys Tyr Ser Asn Val Ile Thr Asn Val Thr Phe Trp Gly Leu Lys
 645 650 655
 Asp Asp Tyr Ser Trp Lys Asn Ala Arg Arg Asn Asp Trp Pro Leu Leu
 660 665 670
 Phe Asp Lys Asp Tyr Gln Ala Lys Leu Ala Tyr Trp Ala Ile Val Ser
 675 680 685
 Pro Glu Ala Leu Pro Val Leu Pro Lys Lys Trp Ser Ile Ala Thr Gly
 690 695 700
 Ser Ala Leu Val Val Gly Met Met Asp Asp Ser Tyr Leu Ala Ser Ser
 705 710 715 720
 Pro Ile Lys Ile Leu Val Asp Gly Gln Glu Lys Leu Thr Ala Arg Val
 725 730 735
 Ile Trp Glu Glu Asn Lys Leu Phe Val Tyr Ala Glu Val Tyr Asp Arg
 740 745 750
 Thr Arg Asp Lys Gly Lys Asp Gly Ile Thr Ile Phe Val Asp Pro Lys
 755 760 765
 Asn Phe Lys Ala Pro Tyr Leu His Glu Asp Ala Phe Tyr Val Thr Ile
 770 775 780
 Lys Thr Asp Trp Ser Val Glu Lys Ser Arg Asp Ile Glu Val Gln
 785 790 795 800
 Arg Phe Val Gly Pro Ser Gly Val Arg Tyr Asn Val Glu Cys Glu Ile
 805 810 815
 Thr Leu Pro Glu Lys Leu Gln Glu Gly Gln Gln Ile Gly Phe Asp Ile
 820 825 830
 Ala Val Gln Asp Gly Asp Lys Val Tyr Ser Trp Ser Asp Thr Ser Asn
 835 840 845
 Gln Gln Lys Leu Ala Thr Met Asn Tyr Gly Thr Leu Thr Leu Gln Gly
 850 855 860
 Ala Val Met Ala Thr Ala Lys Tyr Gly Thr Pro Val Ile Asp Gly Glu
 865 870 875 880
 Ile Asp Asp Ile Trp Tyr Thr Thr Glu Glu Ile Ser Thr Asp Val Val
 885 890 895
 Val Met Gly Ser Leu Lys Asn Ala Arg Ala Lys Val Arg Val Leu Trp
 900 905 910
 Asp Glu Glu His Leu Tyr Val Leu Ala Ile Val Thr Asp Pro Val Leu
 915 920 925
 Asn Lys Asp Asn Thr Asn Pro Trp Glu Gln Asp Ser Val Glu Ile Phe
 930 935 940
 Ile Asp Glu Asn Asn Ala Lys Thr Pro Tyr Tyr Gln Asp Asp Asp Ala
 945 950 955 960
 Gln Tyr Arg Val Asn Tyr Leu Asn Glu Gln Ser Phe Gly Thr Gly Ala
 965 970 975
 Ser Ser Lys Asn Phe Lys Thr Ala Val Lys Leu Ile Asp Gly Gly Tyr
 980 985 990
 Leu Val Glu Ala Ala Val Lys Trp Lys Thr Ile Lys Pro Ser Pro Asn
 995 1000 1005
 Thr Val Ile Gly Phe Asp Phe Gln Val Asn Asp Ala Asn Ala Gln Gly
 1010 1015 1020
 Lys Arg Val Gly Ile Leu Lys Trp Cys Asp Pro Thr Asp Asn Ser Trp
 1025 1030 1035 1040
 Gln Asn Thr Ser Lys Phe Gly Asn Leu Arg Leu Ile Lys
 1045 1050

<210> 53
 <211> 2370
 <212> DNA
 <213> Bacteria

<400> 53
 atgaagggcc tgcaccggct ccgccgccgt cgccggacct gggtggcagg actgtcggcc 60
 gcggcggtgg tcgccggcgc cctgacgctc ctcccggct ccgccggcgc cgcgggcctg 120
 ggtacgcacg cggccccctc gggccggtac ttccggcacg ccgtggccgc gggccgcctc 180
 ggcgactcgg cgtacaccgc gatcgccgac cgggagtcca acatgatcac cccggagaac 240
 gagatgaagt gggacgccgt cgagccgtcc cgcgccggtt tcgacttcgg tcccgcggac 300
 cggatcgtcg agcgtgccct ggcacgcggc cagcgcgtcc gcggccacac cacggtctgg 360
 cactcgagc tcccctcctg ggtgggtcc atccgcgaca cgaagacgct gcgggcgtg 420
 atgaaccacc acatcaccac ccagatgacc cactacaagg gcaagatcta cgctgggac 480

```

gtggtcaacg aggccttcgc cgatggcggc agcggccggc tccgcgactc ggtcttccag 540
aagggtgctg gcgacggctt catcgaggag gcgttccgca ccgcccgcgc ggccgacccc 600
tcggccaagc tctgctacaa cgactacaac atcgagaact ggtcggacgc caagacccag 660
ggcgtctacc gcctggtgaa ggacttcacg tcccagggcg tcccacatga ctgcgtcggc 720
ttccagagcc acttcggcgc gggcgggccc ccggcgagct tcaagacgac cctggccaac 780
ttcgccgccc tgggcgtcga cgtccagatc accgagctgg acatcgcca ggcatcacct 840
gccactacg cgagtgcggt cagcacctgt ctgtccgtgg cccggtgcac cggcatcacg 900
gtgtggggcg tccgtgacag cgactcctgg cggagcgccg aaagcccgt gctgttcgac 960
cggaaaggca agcccaagcc cgcgtacgcc gccgtcatga acgcccctcg ctccggctcg 1020
ggtccacccc cgagcaagcc ggccgacggg acggggagcg gtacggggga gatcaagggc 1080
gtggcctccg gccgctgtct ccaggccaac gcctccacca ccgccaacgg caccggggcg 1140
cagctgtggg actgtagcgg ccaggccaac cagcgtggga cccacaccgc cggcaagcag 1200
ctgaagatcc acggcgacaa gtgcctggac gccaagggca agggcaccgc caacggcacc 1260
gcggtggtcg tctgggactg caacggcgcc gtccggtctg tgcctcgatg ccgtcggcgc ggccaccgcc 1320
ggcacgatca ccggcgctca gtccggtctg gcacgcctgt gggggtgtcg gcaaccagaa gtggtccgcc 1380
aacggcaccc cgatccagct gcacgcctgt gggggtgtcg cgtacaagt gagctcgacg 1440
ccgtccggat cgggcggcgg cacgtgcgtg cttccgtcga cgtacaagt gagctcgacg 1500
ggtgcccctg cgcagcccaa ggccgggtgg gcctcgctga aggacttcac ccatgtgtgtg 1560
ctgggcggca agcacctggt ctacgggtcg aacttcaacg gatcgacgta cggctcgatg 1620
acgttcagcc ccttcaccac ctgtttctac ttccgaccca agaagatctg ggtgctggcg 1680
cagcccggcg tcgcacccac cctgtttctac ttccgaccca agaagatctg ggtgctggcg 1740
taccagtggg gcagggaccgc gttctcttac cggacgtcga ccgacccac caaccgaac 1800
ggctggtcgg cggagcagga gctcttctcc ggaagcatca ccggctcggg cagggcccc 1860
atcgaccaga cgctcatcgg cgacgggacg aacatgtacc tgttcttcgc cggtgacaac 1920
ggcaagatct accgggccag catgcccgat ggggaacttc cgggcagctt cggctcctcg 1980
tacacgacgg tcatgagcga caccgcgaag aacctgttcg aggcgccgca ggtgtacaag 2040
gtcaaggacc agaaccagta cctcatgatc gtcgagggcc ggggcgcggg cgagcgccgc 2100
tacttccgct cgttcacggc ctccagcctg agcgggtgct ggaccccga ggccgcgacc 2160
gagagcaacc ccttcgcggg caaggccaac agcggcgcca cctggaccga cgacatcagc 2220
cacggtgatc tgatcgccac caaccggat cagaccatga ccatcgaccc ctgcaacctt 2280
cagctgctct accagggcaa gtccccgag gcgggcggac cctacgacca gctgccgtac 2340
cggccgggcg tcctcacctt gcagcgctga

```

<210> 54
 <211> 787
 <212> PRT
 <213> Bacteria

<220>
 <221> SIGNAL
 <222> (1)...(37)

<400> 54
 Met Lys Gly Leu His Arg Leu Arg Arg Arg Arg Thr Trp Val Ala
 1 5 10 15
 Gly Leu Ser Ala Ala Ala Val Val Ala Gly Ala Leu Thr Leu Leu Pro
 20 25 30
 Gly Ser Ala Gly Ala Ala Gly Leu Gly Thr His Ala Ala Pro Ser Gly
 35 40 45
 Arg Tyr Phe Gly Thr Ala Val Ala Ala Gly Arg Leu Gly Asp Ser Ala
 50 55 60
 Tyr Thr Ala Ile Ala Asp Arg Glu Phe Asn Met Ile Thr Pro Glu Asn
 65 70 75 80
 Glu Met Lys Trp Asp Ala Val Glu Pro Ser Arg Gly Arg Phe Asp Phe
 85 90 95
 Gly Pro Ala Asp Arg Ile Val Glu Arg Ala Leu Ala Arg Gly Gln Arg
 100 105 110
 Val Arg Gly His Thr Thr Val Trp His Ser Gln Leu Pro Ser Trp Val
 115 120 125
 Gly Ser Ile Arg Asp Thr Lys Thr Leu Arg Gly Val Met Asn His His
 130 135 140
 Ile Thr Thr Gln Met Thr His Tyr Lys Gly Lys Ile Tyr Ala Trp Asp
 145 150 155 160
 Val Val Asn Glu Ala Phe Ala Asp Gly Gly Ser Gly Arg Leu Arg Asp
 165 170 175
 Ser Val Phe Gln Lys Val Leu Gly Asp Gly Phe Ile Glu Glu Ala Phe
 180 185 190
 Arg Thr Ala Arg Ala Ala Asp Pro Ser Ala Lys Leu Cys Tyr Asn Asp
 195 200 205

Tyr Asn Ile Glu Asn Trp Ser Asp Ala Lys Thr Gln Gly Val Tyr Arg
 210 220
 Leu Val Lys Asp Phe Thr Ser Arg Gly Val Pro Ile Asp Cys Val Gly
 225 230 235 240
 Phe Gln Ser His Phe Gly Ala Gly Gly Pro Pro Ala Ser Phe Lys Thr
 245 250 255
 Thr Leu Ala Asn Phe Ala Ala Leu Gly Val Asp Val Gln Ile Thr Glu
 260 265 270
 Leu Asp Ile Ala Gln Ala Ser Pro Ala His Tyr Ala Ser Ala Val Ser
 275 280 285
 Thr Cys Leu Ser Val Ala Arg Cys Thr Gly Ile Thr Val Trp Gly Val
 290 295 300
 Arg Asp Ser Asp Ser Trp Arg Ser Ala Glu Ser Pro Leu Leu Phe Asp
 305 310 315 320
 Arg Asn Gly Lys Pro Lys Pro Ala Tyr Ala Val Met Asn Ala Leu
 325 330 335
 Gly Ser Gly Ser Gly Pro Thr Pro Ser Lys Pro Ala Asp Gly Thr Gly
 340 345 350
 Ser Gly Thr Gly Glu Ile Lys Gly Val Ala Ser Gly Arg Cys Leu Asp
 355 360 365
 Val Pro Ala Ser Thr Thr Ala Asn Gly Thr Arg Ala Gln Leu Trp Asp
 370 375 380
 Cys Ser Gly Gln Ala Asn Gln Arg Trp Thr His Thr Ala Gly Lys Gln
 385 390 395 400
 Leu Lys Ile His Gly Asp Lys Cys Leu Asp Ala Lys Gly Lys Gly Thr
 405 410 415
 Ala Asn Gly Thr Ala Val Val Val Trp Asp Cys Asn Gly Gly Thr Asn
 420 425 430
 Gln Gln Trp Asn Val His Thr Asp Gly Thr Ile Thr Gly Val Gln Ser
 435 440 445
 Gly Leu Cys Leu Asp Ala Val Gly Ala Ala Thr Ala Asn Gly Thr Pro
 450 455 460
 Ile Gln Leu His Ala Cys Gly Gly Val Gly Asn Gln Lys Trp Ser Ala
 465 470 475 480
 Pro Ser Gly Ser Gly Gly Thr Cys Val Leu Pro Ser Thr Tyr Lys
 485 490 495 500
 Trp Ser Ser Thr Gly Ala Leu Ala Gln Pro Lys Ala Gly Trp Ala Ser
 505 510 515
 Leu Lys Asp Phe Thr His Val Val Leu Gly Gly Lys His Leu Val Tyr
 520 525
 Gly Ser Asn Phe Asn Gly Ser Thr Tyr Gly Ser Met Thr Phe Ser Pro
 530 535 540
 Phe Thr Thr Trp Ser Asp Met Ala Ser Ala Gly Gln Lys Ala Met Lys
 545 550 555 560
 Gln Pro Ala Val Ala Pro Thr Leu Phe Tyr Phe Ala Pro Lys Lys Ile
 565 570 575
 Trp Val Leu Ala Tyr Gln Trp Gly Arg Thr Ala Phe Ser Tyr Arg Thr
 580 585 590
 Ser Thr Asp Pro Thr Asn Pro Asn Gly Trp Ser Ala Glu Gln Glu Leu
 595 600 605
 Phe Ser Gly Ser Ile Thr Gly Ser Gly Thr Gly Pro Ile Asp Gln Thr
 610 615 620
 Leu Ile Gly Asp Gly Thr Asn Met Tyr Leu Phe Phe Ala Gly Asp Asn
 625 630 635 640
 Gly Lys Ile Tyr Arg Ala Ser Met Pro Ile Gly Asn Phe Pro Gly Ser
 645 650 655
 Phe Gly Ser Ser Tyr Thr Thr Val Met Ser Asp Thr Ala Lys Asn Leu
 660 665 670
 Phe Glu Ala Pro Gln Val Tyr Lys Val Lys Asp Gln Asn Gln Tyr Leu
 675 680 685
 Met Ile Val Glu Ala Arg Gly Ala Gly Glu Arg Arg Tyr Phe Arg Ser
 690 695 700
 Phe Thr Ala Ser Ser Leu Ser Gly Ala Trp Thr Pro Gln Ala Ala Thr
 705 710 715 720
 Glu Ser Asn Pro Phe Ala Gly Lys Ala Asn Ser Gly Ala Thr Trp Thr
 725 730 735
 Asp Asp Ile Ser His Gly Asp Leu Ile Arg Thr Asn Pro Asp Gln Thr
 740 745 750
 Met Thr Ile Asp Pro Cys Asn Leu Gln Leu Leu Tyr Gln Gly Lys Ser

755
 Pro Gln Ala Gly Gly Pro Tyr Asp Gln Leu Pro Tyr Arg Pro Gly Val
 770 775 780
 Leu Thr Leu
 785

<210> 55
 <211> 1143
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 55
 atgaaaaaaa cgattgcaca tttcacctta tggatagcgt tttttctctt cacttcctgt 60
 gctgttacgg cgcagaagaa tactaagaat gcaagagtaa agcccactac tctaaaagag 120
 gcttaccagg gtaaattcta tatcggtaca gcgatgaatc tgagacagat tcacggagat 180
 gatccccagt ctgaaaatat tatcaaaaaa cagttcaatt ccattgttgc tgaaaactgc 240
 atgaagagta tgtatcttca gccggaggaa ggaaaatttt tcttcgatga tgcggataag 300
 tttgtggatt ttggtcttca gaacaatatg tttattatcg ggcattgtct gatttggcat 360
 tcgcaggcgc caaaatgggt tttcaccgac gagaatggga aaacgggtctc cccagaagtt 420
 cttaaacaaa ggatgaaagc tcatatcacc gccgtcgttt ctcgctacaa agggaaaaatc 480
 aaaggatggg atgtggtgaa cgaagccatt atggaagatg gttcttaccg caaaagcaaa 540
 ttttacgaga ttttgggaga agaatttatt ccgttggcat ttcagtatgc gcatgaagca 600
 gatcctgatg cagaactcta ttacaacgat tataacgaat ggtatcccgg aaaaagagct 660
 acggtgacca aaataatccg agatttcaaa tctagaggaa tccgcattga tgccattgga 720
 atgcaggctc atttcgggat ggattcacc actatagaag agtatgaaca aactattcag 780
 ggctatataa aagaaggcgt gaaagtcaat attacggaac tcgatttgag tccacttcct 840
 tccccctggg gaacttccgc caacgttgcc gatacgcagc agtatcagga aaaaatgaat 900
 ccttacacca aaggacttcc cacagagggt gaaaaagctt gggaaaaccg ttatctcgat 960
 tttttcaaac tattcctaaa atatcatcag catatcgagc gtgttacgtt ttggggcggt 1020
 agcgatatcg attcctggaa gaacgatttt ccagtgagag gacgtaccga ttatccgtta 1080
 ccctttgacc gacagtatca ggcaaaacct ttggttcaga aattaataga cttaacgaaa 1140
 tag 1143

<210> 56
 <211> 380
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(24)

<400> 56
 Met Lys Lys Thr Ile Ala His Phe Thr Leu Trp Ile Ala Phe Phe Leu
 1 5 10 15
 Phe Thr Ser Cys Ala Val Thr Ala Gln Lys Asn Thr Lys Asn Ala Arg
 20 25 30
 Val Lys Pro Thr Thr Leu Lys Glu Ala Tyr Gln Gly Lys Phe Tyr Ile
 35 40 45
 Gly Thr Ala Met Asn Leu Arg Gln Ile His Gly Asp Asp Pro Gln Ser
 50 55 60
 Glu Asn Ile Ile Lys Lys Gln Phe Asn Ser Ile Val Ala Glu Asn Cys
 65 70 75 80
 Met Lys Ser Met Tyr Leu Gln Pro Glu Glu Gly Lys Phe Phe Phe Asp
 85 90 95
 Asp Ala Asp Lys Phe Val Asp Phe Gly Leu Gln Asn Asn Met Phe Ile
 100 105 110
 Ile Gly His Cys Leu Ile Trp His Ser Gln Ala Pro Lys Trp Phe Phe
 115 120 125
 Thr Asp Glu Asn Gly Lys Thr Val Ser Pro Glu Val Leu Lys Gln Arg
 130 135 140
 Met Lys Ala His Ile Thr Ala Val Val Ser Arg Tyr Lys Gly Lys Ile
 145 150 155 160
 Lys Gly Trp Asp Val Val Asn Glu Ala Ile Met Glu Asp Gly Ser Tyr

Arg Lys Ser Lys 165 Phe Tyr Glu Ile Leu 170 Gly Glu Glu Phe Ile 175 Pro Leu
 Ala Phe Gln Tyr 180 Ala His Glu Ala Asp 185 Pro Asp Ala Glu 190 Leu Tyr Tyr
 Asn Asp Tyr Asn Glu Trp Tyr 200 Gly Lys Arg Ala 205 Thr Val Thr Lys
 Ile Ile Arg Asp Phe Lys Ser Arg Gly Ile Arg 220 Ile Asp Ala Ile Gly
 225 Met Gln Ala His Phe 230 Gly Met Asp Ser Pro Thr Ile Glu Glu Tyr Glu
 Gln Thr Ile Gln 245 Gly Tyr Ile Lys Glu 250 Gly Val Lys Val Asn 255 Ile Thr
 Glu Leu Asp Leu Ser Pro Leu Pro 265 Ser Pro Trp Gly Thr 270 Ser Ala Asn
 Val Ala Asp Thr Gln Gln Tyr 280 Glu Lys Met Asn 285 Pro Tyr Thr Lys
 Gly Leu Pro Thr Glu Val Glu Lys Ala Trp Glu Asn Arg Tyr Leu Asp
 305 Phe Phe Lys Leu Phe 310 Leu Lys Tyr His Gln His Ile Glu Arg Val Thr
 Phe Trp Gly Val Ser Asp Ile Asp Ser Trp Lys Asn Asp Phe Pro Val
 340 Arg Gly Arg Thr Asp Tyr Pro Leu Pro Phe Asp Arg Gln Tyr Gln Ala
 355 Lys Pro Leu Val Gln Lys Leu 360 Ile Asp Leu Thr Lys 380
 370

<210> 57
 <211> 1578
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 57
 atgaaaagaa tgatcggttt gctgctggcc atttgcttgg tgatgacgct ggctggggcc 60
 tgggctgcct cggatacgct ggtctatgca tccagtttgc cagcgggcga tgacgactgg 120
 ttgcaaggg gcgcttccc gggtttaccat accacggagg cgacgctgcg gacggaaggc 180
 cggagcgaca actggaattc tccgggacgc tattttgaac tggtgccgga taatgaatat 240
 acgtgagcg tggaggtcta ccaggacgga gcggacagcg cgaacttcat gatttccctg 300
 gaaaaggttg cggatgggat caccggatgg gaaaaccttg tgcggggaac cgtgaaaaag 360
 ggtgaatgga cgacgtgtc cggaacctat acttttgcag actatgaaag ctatgtgctg 420
 tatgtggaga cctccgacgc gccgacgctg gactttgaga tccggaattt ccgggtggaa 480
 agccccaatg ggatcccgga gccgaaggct accgaggcgc cggcagtggt ttcggaagcc 540
 acggatattc cgagcctgaa ggacgcttac gcggattact tgcactttgg cgcggccgtg 600
 ccgcagtctg ctttcaccag cagagataat attcagctga tggagctgat gaaaaaccag 660
 ttcagcatcc tgacgcctga aaatgagctg aagccggaca gtgtattgga tgaagcgcc 720
 agcaagcagc tggccaaaga ggatgaaacc gcggtagtgg tgcggtttaa cggggcaaag 780
 tcattgctgc ggtttgccca gcaaaacggc atcaaggctg acgggcatgt gctggctctg 840
 cacagccaga cgccggaagc ctttttccat gaaggatatg atcccaagaa cccgctggtg 900
 agccgggaag tgatgctggg acggctggaa aactatatcc gggaagtgtc gacccagacg 960
 gaagaactgt atccggcgt gatcgtcagc tgggacgtgg tgaacgaagc gattgacgac 1020
 ggaaccaact ggatccggaa gggatcgggc tggatccgga ccatcgggga agactatgtg 1080
 gagaaggctt ttgagtttgc ccggaagtat gccccggaag gcgtgtgtgct gtactacaac 1140
 gattacaaca cggcatacgc cggaaaactg aatgggatta tcaaactgat caaaccatg 1200
 atcgagcagg gaacgatcga cggatacggc ttccagatgc accatacgac cgggcagccc 1260
 agcaaccaga tgatcaccac ggcggtggag aagatcgcgg ccctgggaat caagctgcgg 1320
 gtcagcgaga tggacatcgg gattacaaag tatacagaga cgagcctgca ggcacaaaag 1380
 gacaagtaca aggcgatgat ggaactgatg ctgcggttcg cggaccagac ggaagcagtg 1440
 caggtctggg ggattacgga tacgatgagc tggcggagct ccagctatcc gctgctgtt 1500
 gaccgagca ggaatccgaa gccgcggttc tatggcgtga ttgaagcggg tgaagactgg 1560
 acagggaaaa gtgaatag

<210> 58
 <211> 525
 <212> PRT
 <213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(22)

<400> 58

```

Met Lys Arg Met Ile Gly Leu Leu Leu Ala Ile Cys Leu Val Met Thr
 1      5      10      15
Leu Ala Gly Ala Trp Ala Ala Ser Asp Thr Leu Val Tyr Ala Ser Ser
 20      25      30
Phe Ala Ala Gly Asp Asp Asp Trp Phe Ala Arg Gly Ala Ser Arg Val
 35      40      45
Tyr His Thr Thr Glu Ala Thr Leu Arg Thr Glu Gly Arg Ser Asp Asn
 50      55      60
Trp Asn Ser Pro Gly Arg Tyr Phe Glu Leu Val Pro Asp Asn Glu Tyr
 65      70      75      80
Thr Leu Ser Val Glu Val Tyr Gln Asp Gly Ala Asp Ser Ala Asn Phe
 85      90      95
Met Ile Ser Leu Glu Lys Val Ala Asp Gly Ile Thr Gly Trp Glu Asn
100      105      110
Leu Val Arg Gly Thr Val Lys Lys Gly Glu Trp Thr Thr Leu Ser Gly
115      120      125
Thr Tyr Thr Phe Ala Asp Tyr Glu Ser Tyr Val Leu Tyr Val Glu Thr
130      135      140
Ser Asp Ala Pro Thr Leu Asp Phe Glu Ile Arg Asn Phe Arg Val Glu
145      150      155      160
Ser Pro Asn Gly Ile Pro Glu Pro Lys Ala Thr Glu Ala Pro Ala Val
165      170      175
Val Ser Glu Ala Thr Asp Ile Pro Ser Leu Lys Asp Ala Tyr Ala Asp
180      185      190
Tyr Phe Asp Phe Gly Ala Ala Val Pro Gln Ser Ala Phe Thr Ser Arg
195      200      205
Asp Asn Ile Gln Leu Met Glu Leu Met Lys Asn Gln Phe Ser Ile Leu
210      215      220
Thr Pro Glu Asn Glu Leu Lys Pro Asp Ser Val Leu Asp Val Ser Ala
225      230      235      240
Ser Lys Gln Leu Ala Lys Glu Asp Glu Thr Ala Val Val Val Arg Phe
245      250      255
Asn Gly Ala Lys Ser Leu Leu Arg Phe Ala Gln Gln Asn Gly Ile Lys
260      265      270
Val His Gly His Val Leu Val Trp His Ser Gln Thr Pro Glu Ala Phe
275      280      285
Phe His Glu Gly Tyr Asp Pro Lys Asn Pro Leu Val Ser Arg Glu Val
290      295      300
Met Leu Gly Arg Leu Glu Asn Tyr Ile Arg Glu Val Leu Thr Gln Thr
305      310      315      320
Glu Glu Leu Tyr Pro Gly Val Ile Val Ser Trp Asp Val Val Asn Glu
325      330      335
Ala Ile Asp Asp Gly Thr Asn Trp Ile Arg Lys Gly Ser Gly Trp Tyr
340      345      350
Arg Thr Ile Gly Glu Asp Tyr Val Glu Lys Ala Phe Glu Phe Ala Arg
355      360      365
Lys Tyr Ala Pro Glu Gly Val Leu Leu Tyr Tyr Asn Asp Tyr Asn Thr
370      375      380
Ala Tyr Ala Gly Lys Leu Asn Gly Ile Ile Lys Leu Ile Lys Pro Met
385      390      395      400
Ile Glu Gln Gly Thr Ile Asp Gly Tyr Gly Phe Gln Met His His Thr
405      410      415
Thr Gly Gln Pro Ser Asn Gln Met Ile Thr Thr Ala Val Glu Lys Ile
420      425      430
Ala Ala Leu Gly Ile Lys Leu Arg Val Ser Glu Met Asp Ile Gly Ile
435      440      445
Thr Lys Tyr Thr Glu Thr Ser Leu Gln Ala Gln Lys Asp Lys Tyr Lys
450      455      460
Ala Met Met Glu Leu Met Leu Arg Phe Ala Asp Gln Thr Glu Ala Val
465      470      475      480
Gln Val Trp Gly Ile Thr Asp Thr Met Ser Trp Arg Ser Ser Ser Tyr

```

485 490 495
 Pro Leu Leu Phe Asp Arg Ser Arg Asn Pro Lys Pro Ala Phe Tyr Gly
 500 505 510
 Val Ile Glu Ala Val Glu Asp Trp Thr Gly Lys Ser Glu
 515 520 525

<210> 59
 <211> 1104
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 59
 atgcttgcca gtagtgccgg tttggtagca tcccaactca agctgtccgc gtttagctgca 60
 gctaaaaaatg ctggattaaa agatgtatat aaggatcgct ttctgattgg tgcagcaatt 120
 aatacctcga ttgcgagcgg ccagcaacct gatattacag aaattatcaa gcgtgatttt 180
 tcgctcgttaa cacctgaaaa tgcaatgaag tgggaatctg tcaggactgc tgatggcggg 240
 tggaaatggg cagatgccga tcaattcgtt acgtttgcaa cagaacacaa aatacacgct 300
 gttggccaca cccttgccctg gcatagccag attcccgaatt ccgtattcaa aaatgaaaaa 360
 ggcgaaataca taaaatccac cgagctatca aaaaaaatgg aagaacatat cactacgatt 420
 gtaggtagat ataaaggcaa actcgatgcc tgggatgtag ttaatgaggc tgttggtgat 480
 gataatcaaa tgcgcaaaaag ccattattac aatattctcg gcgaagattt tattgataag 540
 gcatttcacc ttgcgcatga ggtcgatccc aaagcgcatt taatgtataa cgactacaac 600
 attgaaaaag atggcaagcg tgaagctacc cttgaaatgt taaagcgttt acaaaaacgc 660
 ggtgtaccga ttcatgggct cggcatccag ggacatattg ccgttgatgg ccccagcatt 720
 gcggatattg aaaaaagtat tttggcttat gcggatttgg gtttgcggtg acatttcacc 780
 gagttggata ttgatgtatt gccgcaaatc tggaaacttac cggttgcaga aatttctaca 840
 cgcttcgaat acaaacctga gcgagatcct ttcaaaaatg gtttatcaaa agaaatgaac 900
 gataaactca gtgcacgcta tgaagaatta ttcacattat ttattaaaca caaagataaa 960
 attgatcgta ttactttgtg ggggtgtcagc gatgatgcaa cctggctaaa tgatttcccc 1020
 atcaaaggca gaaccagtta tccattattg tttgatcgca agcatcaacc aaaagatgct 1080
 tattataaca ttctggcggt gtga 1104

<210> 60
 <211> 367
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(21)

<400> 60
 Met Leu Ala Ser Ser Ala Gly Leu Val Ala Ser Gln Leu Lys Leu Ser
 1 5 10 15
 Ala Leu Ala Ala Ala Lys Asn Ala Gly Leu Lys Asp Val Tyr Lys Asp
 20 25 30
 Arg Phe Leu Ile Gly Ala Ala Ile Asn Thr Ser Ile Ala Ser Gly Gln
 35 40 45
 Gln Pro Asp Ile Thr Glu Ile Ile Lys Arg Asp Phe Ser Ser Leu Thr
 50 55 60
 Pro Glu Asn Ala Met Lys Trp Glu Ser Val Arg Thr Ala Asp Gly Gly
 65 70 75 80
 Trp Lys Trp Ala Asp Ala Asp Gln Phe Val Thr Phe Ala Thr Glu His
 85 90 95
 Lys Ile His Ala Val Gly His Thr Leu Ala Trp His Ser Gln Ile Pro
 100 105 110
 Asp Ser Val Phe Lys Asn Glu Lys Gly Glu Tyr Ile Lys Ser Thr Glu
 115 120 125
 Leu Ser Lys Lys Met Glu Glu His Ile Thr Thr Ile Val Gly Arg Tyr
 130 135 140
 Lys Gly Lys Leu Asp Ala Trp Asp Val Val Asn Glu Ala Val Gly Asp
 145 150 155 160
 Asp Asn Gln Met Arg Lys Ser His Tyr Tyr Asn Ile Leu Gly Glu Asp
 165 170 175

Phe Ile Asp Lys Ala Phe His Leu Ala His Glu Val Asp Pro Lys Ala
 180 185 190
 His Leu Met Tyr Asn Asp Tyr Asn Ile Glu Lys Asp Gly Lys Arg Glu
 195 200 205
 Ala Thr Leu Glu Met Leu Lys Arg Leu Gln Lys Arg Gly Val Pro Ile
 210 215 220
 His Gly Leu Gly Ile Gln Gly His Ile Ala Val Asp Gly Pro Ser Ile
 225 230 235 240
 Ala Asp Ile Glu Lys Ser Ile Leu Ala Tyr Ala Asp Leu Gly Leu Arg
 245 250 255
 Val His Phe Thr Glu Leu Asp Ile Asp Val Leu Pro Gln Ile Trp Asn
 260 265 270
 Leu Pro Val Ala Glu Ile Ser Thr Arg Phe Glu Tyr Lys Pro Glu Arg
 275 280 285
 Asp Pro Phe Lys Asn Gly Leu Ser Lys Glu Met Asn Asp Lys Leu Ser
 290 295 300
 Ala Arg Tyr Glu Glu Leu Phe Thr Leu Phe Ile Lys His Lys Asp Lys
 305 310 315 320
 Ile Asp Arg Ile Thr Leu Trp Gly Val Ser Asp Asp Ala Thr Trp Leu
 325 330 335
 Asn Asp Phe Pro Ile Lys Gly Arg Thr Ser Tyr Pro Leu Leu Phe Asp
 340 345 350
 Arg Lys His Gln Pro Lys Asp Ala Tyr Tyr Asn Ile Leu Ala Leu
 355 360 365

<210> 61
 <211> 1041
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 61
 atgagaagaa gcatggaaag gctgccaag ctccatgaag cttacggcaa tagtttcaag 60
 atcggcgctg ccgtgaatcc aattacgatg gtgacccaaa aggaattgtt gtcacaccac 120
 ttcaacagcg ttacggcaga aaatgaaatg aaattcgagc gattgcaccc atcgggaagag 180
 gtgtatacat tcgagcaagc cgaccagatc gtatcgtttg ccaaatacgaa cggaatgtcg 240
 gtgagaggac ataccctcgt atggcataat cagacgccgg aatgggtgtt tcaagacagt 300
 tccggtggga cagccggccg cgagctgctg ctcgctcgga tgaaatcgca catcgatgag 360
 gtcgttgccc gttatcgccg agatatctat gcttgggatg tcgtaaacga agccattgcc 420
 gacagtggaa gcgatctgct tcgttcctcc ccgtggcttg cgtcgatcgg ggaggatttt 480
 atcgccaagg ctttcgaata tgcgcacgaa gcagaccgcg aagcgctgct gttttataac 540
 gattacaacg aatccgtgcc cgagaagcgg gagaagattt acacgctcct taaatcgta 600
 aaggagcagg atgtgccgat tcacggcgct gggcttcagg cccattggaa tttggagt 660
 ccatcgcttg acgatatccg cagggcaatc gaaaggtatg caagccttgg catgatcttg 720
 catatcacgg agcttgacgt atccgtattc gcgcatgagg ataagcggac cgaatctggc 780
 gcgccgaccg acgaaatgct tgagcgccag gcggagcgtt acggtcaatt gttccgtctg 840
 ctaaaagagt acagcggcag gtcacttcc gtgaccttct ggggagcggc ggacgattat 900
 acctggctgg atcattttcc ggtaaggggc cgcaaaaatt ggccgttcgt cttcgacgag 960
 aacctcttc cgaaggaatc ctattggaac ctgttggaag aagccaatcc cgaaagaaca 1020
 ttccaagaga tacgttcgta a 1041

<210> 62
 <211> 346
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 62
 Met Arg Arg Ser Met Glu Arg Leu Pro Lys Leu His Glu Ala Tyr Gly
 1 5 10 15
 Asn Ser Phe Lys Ile Gly Ala Ala Val Asn Pro Ile Thr Met Val Thr
 20 25 30
 Gln Lys Glu Leu Leu Ser His His Asn Ser Val Thr Ala Glu Asn
 35 40 45
 Glu Met Lys Phe Glu Arg Leu His Pro Ser Glu Glu Val Tyr Thr Phe

<210>	64
<211>	369

<212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(20)

<400> 64
 Met Lys Arg Ile Leu Ile Gly Leu Ala Ala Leu Thr Ala Ser Gly Leu
 1 5 10 15
 Ser Ala Gln Lys Ser Asp Gly Thr Leu Lys Lys Ala Phe Gln Asp Lys
 20 25 30
 Phe Tyr Ile Gly Thr Ala Met Ser Leu Pro Gln Ile Asp Gly Thr Asp
 35 40 45
 Lys Arg Ala Val Ala Ile Ile Arg Asn Gln Phe Ser Ile Val Ala
 50 55 60
 Glu Asn Cys Met Lys Ser Met Phe Leu Gln Pro Gln Glu Gly Lys Phe
 65 70 75 80
 Phe Phe Asp Asp Ala Asp Lys Phe Val Asp Phe Gly Met Lys Asn Asn
 85 90 95
 Met Phe Val Ile Gly His Thr Leu Ile Trp His Ser Gln Leu Pro Lys
 100 105 110
 Trp Phe Phe Thr Asp Lys Asn Gly Lys Asp Val Ser Pro Glu Val Leu
 115 120 125
 Lys Gln Arg Met Lys Asn His Ile Thr Thr Val Val Ser Arg Tyr Lys
 130 135 140
 Gly Lys Val Lys Gly Trp Asp Val Val Asn Glu Ala Ile Leu Glu Asp
 145 150 155 160
 Gly Thr Tyr Arg Lys Ser Lys Phe Tyr Glu Ile Leu Gly Glu Asp Phe
 165 170 175
 Ile Pro Leu Ala Phe Gln Tyr Ala Gln Glu Ala Asp Pro Asn Ala Glu
 180 185 190
 Leu Tyr Tyr Asn Asp Tyr Asn Glu Trp Tyr Pro Glu Lys Val Lys Ala
 195 200 205
 Val Ile Thr Met Val Glu Lys Leu Lys Ser Arg Gly Ile Arg Ile Asp
 210 215 220
 Gly Val Gly Met Gln Ala His Val Gly Met Asp Ile Pro Ser Ile Asn
 225 230 235 240
 Glu Tyr Glu Lys Ala Ile Leu Ala Tyr Ser Asn Ala Gly Val Lys Val
 245 250 255
 Asn Ile Thr Glu Leu Glu Ile Ser Ala Leu Pro Ser Pro Trp Gly Ser
 260 265 270
 Ser Ala Asn Val Ser Asp Thr Val Ala Tyr Gln Lys Glu Met Asn Pro
 275 280 285
 Tyr Thr Lys Gly Leu Pro Asn Glu Val Glu Ala Lys Trp Glu Lys Arg
 290 295 300
 Tyr Leu Asp Phe Phe Ser Leu Phe Leu Lys His Lys Asp Lys Ile Arg
 305 310 315 320
 Arg Val Thr Leu Trp Gly Val Thr Asp Lys Gln Ser Trp Lys Asn Asp
 325 330 335
 Phe Pro Val Lys Gly Arg Thr Asp Tyr Pro Leu Leu Phe Asp Arg Lys
 340 345 350
 Asp Gln Glu Lys Pro Val Val Gln Lys Ile Ile Lys Leu Ala Glu Lys
 355 360 365
 Asn

<210> 65
 <211> 1557
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 65
 atgaaaagaa tcggactgtt gctgctggct gtgatcatgc ttgtgggctg tgtatattcc
 Page 51

gcggcgccg	aggatacgt	ggtttatgt	tccacttttg	tggccggaac	ggacggatgg	120
tacgcccgcg	gagcgagaa	agtataccgc	acaaccgag	agacactgcg	gacggaaggc	180
cggaccagcg	actggcattc	cccgggcccgt	gattttgacc	tgggtggaagg	cggcggtctat	240
gtcctgagcg	tggaagtgtt	ccaggacgaa	gcggacaacg	ccagcttcat	gatttccatc	300
gcccacagca	aggacggtac	ggaaacctat	gaaaaccttg	ctcgcggaac	cgccaaacgc	360
ggcgagtggg	tcacctgac	cggaaacatat	accgccggca	attttgaccg	gaacgtcctg	420
tatgtggaaa	cgaccggatc	gccggaactg	agctatgaaa	tccggaattt	ccgggttgaa	480
gcgccaagcg	gagttccgga	gccgaaggct	acggagcccc	cgatggtgat	tgaggcgggtg	540
gagaacctcc	cgggcctgaa	gaacgcgtat	gcgggaaaat	ttgatttcgg	cgcggcgggtt	600
ccgggatacg	ctttcggcga	tccgggcctg	aaacagctga	tgactgagca	gttcagcatc	660
ctgacgcccc	aaaacgaact	gaaaccggac	gctgtgctgg	acgtggcggc	gagcaagcgg	720
ctggcccagg	aggatgaaac	ggcgggtggc	gttcattttg	acggcgccat	tccgctgctg	780
aactttgccc	gggacaacgg	catcaggggtg	cacggacatg	tgctgatctg	gcacagccag	840
acgccggaag	cgttcttcca	tgagggttat	gacacctcca	agcccctggt	cagccgggaa	900
gtgatgctgg	gccggatgga	aaactatatc	cgcgagggtgc	tgacctggac	gaacgagaat	960
tatccgggcg	tgatcgtatc	ctgggacgtg	gtgaacgaag	ccattgatga	cggaaacgaac	1020
tggctgcgga	attccaactg	gtacaagacg	gtgggaggcg	actttgtgaa	ccgggctttt	1080
gaatttgccc	gcatgtacgc	ggcggacggc	gtcctcctgt	attacaatga	ttacaataacc	1140
gcctatccgg	ccaaacggaa	gggaatcatc	aagctgctgg	gccagctgat	tgaggaaggc	1200
aatattgacg	gatacggctt	ccagatgcat	cacagcaccg	gcgagccttc	catggagatg	1260
atcacgcgtt	cggtggagga	aatcgcccg	ctgggaataa	aactgcgggt	cagcgagctg	1320
gatgtgggca	tgggcagcag	catgacggaa	gaagccctga	tgaacagaa	ggacaaatac	1380
aaggcggtca	tggaactgat	gctgcggttt	gccgaccaga	cggaaagcgg	gcaggtatgg	1440
ggactgacgg	acaatatgag	ctggcggacc	ggccagaatc	cgctgctgtt	tgaccggaac	1500
cggaaaccca	agccggcctt	cttcggcgctc	ctggaagcgg	cggaagaaag	caaataa	1557

<210> 66
 <211> 518
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(22)

<400> 66
 Met Lys Arg Ile Gly Leu Leu Leu Leu Ala Val Ile Met Leu Val Gly
 1 5 10 15
 Cys Val Tyr Ser Ala Ala Ala Glu Asp Thr Leu Val Tyr Ala Ser Thr
 20 25 30
 Phe Val Ala Gly Thr Asp Gly Trp Tyr Ala Arg Gly Ala Gln Lys Val
 35 40 45
 Tyr Arg Thr Thr Glu Glu Thr Leu Arg Thr Glu Gly Arg Thr Ser Asp
 50 55 60
 Trp His Ser Pro Gly Arg Asp Phe Asp Leu Val Glu Gly Gly Val Tyr
 65 70 75 80
 Val Leu Ser Val Glu Val Phe Gln Asp Glu Ala Asp Asn Ala Ser Phe
 85 90 95
 Met Ile Ser Ile Ala His Ser Lys Asp Gly Thr Glu Thr Tyr Glu Asn
 100 105 110
 Leu Ala Arg Gly Thr Ala Lys Arg Gly Glu Trp Val Thr Leu Thr Gly
 115 120 125
 Thr Tyr Thr Ala Gly Asn Phe Asp Arg Asn Val Leu Tyr Val Glu Thr
 130 135 140
 Thr Gly Ser Pro Glu Leu Ser Tyr Glu Ile Arg Asn Phe Arg Val Glu
 145 150 155 160
 Ala Pro Asn Gly Val Pro Glu Pro Lys Ala Thr Glu Pro Pro Met Val
 165 170 175
 Ile Glu Ala Val Glu Asn Leu Pro Gly Leu Lys Asn Ala Tyr Ala Gly
 180 185 190
 Lys Phe Asp Phe Gly Ala Ala Val Pro Gly Tyr Ala Phe Gly Asp Pro
 195 200 205
 Gly Leu Lys Gln Leu Met Thr Glu Gln Phe Ser Ile Leu Thr Pro Glu
 210 215 220
 Asn Glu Leu Lys Pro Asp Ala Val Leu Asp Val Ala Ala Ser Lys Arg
 225 230 235 240
 Leu Ala Gln Glu Asp Glu Thr Ala Val Ala Val His Phe Asp Gly Ala

Ile Pro Leu 245
 260
 His Val Leu Ile Trp His Ser Gln Thr Pro Glu Ala Phe Phe His Glu
 275
 Gly Tyr Asp Thr Ser Lys Pro 280
 290
 Arg Met Glu Asn Tyr Ile Arg Glu Val Leu Thr Trp Thr Asn Glu Asn
 305
 Tyr Pro Gly Val Ile Val Ser Trp Asp Val Val Asn Glu Ala Ile Asp
 310
 325
 Asp Gly Thr Asn Trp Leu Arg Asn Ser Asn Trp Tyr Lys Thr Val Gly
 340
 345
 Gly Asp Phe Val Asn Arg Ala Phe Glu Phe Ala Arg Met Tyr Ala Ala
 355
 360
 Asp Gly Val Leu Leu Tyr Tyr Asn Asp Tyr Asn Thr Ala Tyr Pro Ala
 370
 375
 Lys Arg Lys Gly Ile Ile Lys Leu Leu Gly Gln Leu Ile Glu Glu Gly
 385
 390
 Asn Ile Asp Gly Tyr Gly Phe Gln Met His His Ser Thr Gly Glu Pro
 405
 410
 Ser Met Glu Met Ile Thr Ala Ser Val Glu Glu Ile Ala Ala Leu Gly
 420
 425
 Ile Lys Leu Arg Val Ser Glu Leu Asp Val Gly Met Gly Ser Ser Met
 435
 440
 Thr Glu Glu Ala Leu Met Lys Gln Lys Asp Lys Tyr Lys Ala Val Met
 450
 455
 Glu Leu Met Leu Arg Phe Ala Asp Gln Thr Glu Ala Val Gln Val Trp
 465
 470
 Gly Leu Thr Asp Asn Met Ser Trp Arg Thr Gly Gln Asn Pro Leu Leu
 485
 490
 Phe Asp Arg Asn Arg Asn Pro Lys Pro Ala Phe Phe Gly Val Leu Glu
 500
 505
 Ala Ala Glu Glu Ser Lys
 515

<210> 67
 <211> 1224
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 67
 atgcggaacg tcgtgcgtaa accattgaca atcggactcg cttaacact attattgccc 60
 atgggaatga cggcaacatc agcgaagaat gcagattcct atgcgaaaaa acctcacatc 120
 agcgcatatga atgccccaca attggatcaa cgctacaaaa acgagttcac gattggtgcg 180
 gcagtagaac cttatcaact acaaaatgaa aaagacgtac aaatgctaaa gcgccattc 240
 aacagcattg ttgccgagaa cgtaatgaaa ccgatcagca ttcaacctga ggaaggaaaa 300
 ttcaattttg aacaagcggg tcgaattgtg aagttcgtca aggcgaatgg catggatatt 360
 cgcttcata cactcgtttg gcacagccaa gtacctcaat gggtctttct tgacaaggaa 420
 ggcaagccaa tgggtaatga aacagatcca gtgaaacgtg aacaaaataa acaactgctg 480
 ttaaaacgac ttgaaactca tattaaaacg atcgtcgcgc ggtacaaaga tgacattaag 540
 tactgggacg ttgtaaatga ggttggtggg gacgacggaa aactgcgcaa ctctccatgg 600
 tatcaaatcg ccggcatcga ttatatataa gtggcattcc aaacagcgag aaaatatggc 660
 ggcaacaaga ttaaacttta tatcaatgat tacaataccg aagtggaacc aaagcgaagc 720
 gctctttata acttggtgaa gcaattaaaa gaagagggcg ttctattga cggcatcggc 780
 catcaatccc acattcaaat cggctggcct tctgaagcag aaatcgagaa aacgattaac 840
 atgttcgccc ctctcggtt agacaaccaa atcactgagc ttgatgtgag catgtacggt 900
 tggccgcccgc gcgcttacc gacgtatgac gccattccaa aacaaaagtt tttggatcag 960
 gcagcgcgct atgatcggtt gttcaaacgt tatgaaaagt tgagcgataa aattagcaac 1020
 gtcaccttct ggggcatcgc cgacaatcat acgtggctcg acagccgtgc ggatgtgtac 1080
 tatgacgcca acgggaatgt tgtgggtgac ccgaacgctc cgtacgcaaa agtggaaaaa 1140
 gggaaaggaa aagatgcgcc gttcggtttt ggaccggatt acaaagtcaa acccgcatat 1200
 tgggctatta tcgaccacaa atag 1224

<210> 68
 <211> 407

<212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 68
 Met Arg Asn Val Val Arg Lys Pro Leu Thr Ile Gly Leu Ala Leu Thr
 1 5 10 15
 Leu Leu Leu Pro Met Gly Met Thr Ala Thr Ser Ala Lys Asn Ala Asp
 20 25 30
 Ser Tyr Ala Lys Lys Pro His Ile Ser Ala Leu Asn Ala Pro Gln Leu
 35 40 45
 Asp Gln Arg Tyr Lys Asn Glu Phe Thr Ile Gly Ala Ala Val Glu Pro
 50 55 60
 Tyr Gln Leu Gln Asn Glu Lys Asp Val Gln Met Leu Lys Arg His Phe
 65 70 75 80
 Asn Ser Ile Val Ala Glu Asn Val Met Lys Pro Ile Ser Ile Gln Pro
 85 90 95
 Glu Glu Gly Lys Phe Asn Phe Glu Gln Ala Asp Arg Ile Val Lys Phe
 100 105 110
 Ala Lys Ala Asn Gly Met Asp Ile Arg Phe His Thr Leu Val Trp His
 115 120 125
 Ser Gln Val Pro Gln Trp Phe Phe Leu Asp Lys Glu Gly Lys Pro Met
 130 135 140
 Val Asn Glu Thr Asp Pro Val Lys Arg Glu Gln Asn Lys Gln Leu Leu
 145 150 155 160
 Leu Lys Arg Leu Glu Thr His Ile Lys Thr Ile Val Glu Arg Tyr Lys
 165 170 175
 Asp Asp Ile Lys Tyr Trp Asp Val Val Asn Glu Val Val Gly Asp Asp
 180 185 190
 Gly Lys Leu Arg Asn Ser Pro Trp Tyr Gln Ile Ala Gly Ile Asp Tyr
 195 200 205
 Ile Lys Val Ala Phe Gln Thr Ala Arg Lys Tyr Gly Gly Asn Lys Ile
 210 215 220
 Lys Leu Tyr Ile Asn Asp Tyr Asn Thr Glu Val Glu Pro Lys Arg Ser
 225 230 235 240
 Ala Leu Tyr Asn Leu Val Lys Gln Leu Lys Glu Glu Gly Val Pro Ile
 245 250 255
 Asp Gly Ile Gly His Gln Ser His Ile Gln Ile Gly Trp Pro Ser Glu
 260 265 270
 Ala Glu Ile Glu Lys Thr Ile Asn Met Phe Ala Ala Leu Gly Leu Asp
 275 280 285
 Asn Gln Ile Thr Glu Leu Asp Val Ser Met Tyr Gly Trp Pro Pro Arg
 290 295 300
 Ala Tyr Pro Thr Tyr Asp Ala Ile Pro Lys Gln Lys Phe Leu Asp Gln
 305 310 315 320
 Ala Ala Arg Tyr Asp Arg Leu Phe Lys Leu Tyr Glu Lys Leu Ser Asp
 325 330 335
 Lys Ile Ser Asn Val Thr Phe Trp Gly Ile Ala Asp Asn His Thr Trp
 340 345 350
 Leu Asp Ser Arg Ala Asp Val Tyr Tyr Asp Ala Asn Gly Asn Val Val
 355 360 365
 Val Asp Pro Asn Ala Pro Tyr Ala Lys Val Glu Lys Gly Lys Gly Lys
 370 375 380
 Asp Ala Pro Phe Val Phe Gly Pro Asp Tyr Lys Val Lys Pro Ala Tyr
 385 390 395 400
 Trp Ala Ile Ile Asp His Lys
 405

<210> 69
 <211> 1596
 <212> DNA
 <213> Unknown

<220>

<223> obtained from an environmental sample

<400> 69

atggcgatgc	atagatttaa	gcaattaggg	gccatcctac	ttgtcctatg	gttttgtgca	60
ttgccagtgc	aggcgagggc	ttggcgtgcg	gccgcagagc	agcgtattga	acagtaccgt	120
aagggggccac	tgcgggttca	ggtgaaggat	cctgaaggac	ggcccgtacc	gaatgcccaa	180
gtgcacgttc	gcatgacgcg	tcacgctttt	ggatttggta	cggctgtcag	ctttggcctg	240
gtcgtgggggt	cgggatacaa	ccccacctat	cgggccaaagc	tagaagacct	gacgggagac	300
ggccgcacat	tcaacatggc	tacgccagag	aatgaattga	agtggcctgc	gtgggagtcg	360
gaatggccca	tttcgaatcg	tcgaaagatc	gacgtcatca	actggctgcg	cgcaaaaggc	420
tacagcattc	gaggacacaa	cctgctatgg	cctgactggc	aatggatgcc	ccgtgatatt	480
gagcaaaacc	gcaacaatcc	acagtacatc	tacgatcgcg	ttcgcaatca	cattgcggcg	540
ttggctgggc	atcggggacat	tcggggcaaa	ctgcgggact	gggatgttct	taacgaacca	600
gcccacctga	ccgcatttgcg	cgatgtgttt	aacggttggg	gctcatatga	gcgtggggaa	660
gacttctatg	tggatgtctt	taggtggggc	aaggcagcag	actcgaccgc	ccgtctatac	720
atcaacagat	acaacattat	caacaactac	gccaacgagc	agcctacgcg	caactattac	780
aagtggatca	ttgcacgcct	aatctcaaaa	ggagcgcccta	tcgaagggat	cggcattcag	840
gggcatattt	cggcaccact	gccaagcatg	agtgaagtca	aggcagccct	agacgaaatg	900
gcagtttttg	gattgccttt	ggccatcaca	gaatacgacg	ttaccggcgt	ttcgggaagaa	960
gtcgaagcca	actttatgcg	ggactttttg	accatggctt	ttagtcattc	cgctgtggag	1020
agcttcgtca	tgtgggggtt	ctggagcgga	gcacactggc	gtgacaatgc	gccgctgttt	1080
cgggcccact	ggagtcctca	gccttcggga	caggtgttcc	ttgatctggt	ctttcggcgc	1140
tgggtggaccg	atactacggg	ggtaaccggt	ccagatggca	gctggctctg	acgcggattt	1200
ttaggggatt	acgttgtgga	agtgcaggtg	ggggagggtt	cagtgaacaa	gtccctgcgc	1260
ctcgaagacc	cgcaggatca	aaccacgcta	gaggtggtgg	tcagtagcgt	taaggtgggt	1320
gaaaagcccta	cagaagacgt	gttgcgcgtg	caagggtttg	gaccagaccc	ctttgtcgaa	1380
ggaacggcgc	tgcgctactg	gttagggcgg	ccggccgatg	ttgaactggc	agtgtatgat	1440
gtgctggggc	gacaggtcta	cgccgtgcaa	aagcatcgcg	tagctgggtg	gcatactgaa	1500
tgggtcgagg	cttcccactg	gcctgcagga	ctttatctgt	accgactcca	agcaggtgat	1560
ctgttgcaca	cgggtagaat	ggtcaagatc	caataa			1596

<210> 70

<211> 531

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(25)

<400> 70

Met	Ala	Met	His	Arg	Phe	Lys	Gln	Leu	Gly	Ala	Ile	Leu	Leu	Val	Leu
1				5					10					15	
Trp	Phe	Cys	Ala	Leu	Pro	Val	Gln	Ala	Gln	Ala	Trp	Arg	Ala	Ala	Ala
		20					25					30			
Glu	Gln	Arg	Ile	Glu	Gln	Tyr	Arg	Lys	Gly	Pro	Leu	Arg	Val	Gln	Val
		35					40					45			
Lys	Asp	Pro	Glu	Gly	Arg	Pro	Val	Pro	Asn	Ala	Gln	Val	His	Val	Arg
	50				55				60						
Met	Thr	Arg	His	Ala	Phe	Gly	Phe	Gly	Thr	Ala	Val	Ser	Phe	Gly	Leu
65				70				75						80	
Val	Val	Gly	Ser	Gly	Tyr	Asn	Pro	Thr	Tyr	Arg	Ala	Lys	Leu	Glu	Asp
		85						90						95	
Leu	Thr	Gly	Asp	Gly	Arg	Thr	Phe	Asn	Met	Ala	Thr	Pro	Glu	Asn	Glu
		100					105					110			
Leu	Lys	Trp	Pro	Ala	Trp	Glu	Ser	Glu	Trp	Pro	Ile	Ser	Asn	Arg	Arg
	115					120					125				
Lys	Ile	Asp	Val	Ile	Asn	Trp	Leu	Arg	Ala	Lys	Gly	Tyr	Ser	Ile	Arg
	130				135					140					
Gly	His	Asn	Leu	Leu	Trp	Pro	Asp	Trp	Gln	Trp	Met	Pro	Arg	Asp	Ile
145				150				155						160	
Glu	Gln	Asn	Arg	Asn	Asn	Pro	Gln	Tyr	Ile	Tyr	Asp	Arg	Val	Arg	Asn
		165					170							175	
His	Ile	Ala	Ala	Leu	Ala	Gly	His	Arg	Asp	Ile	Arg	Gly	Lys	Leu	Arg
	180						185					190			
Asp	Trp	Asp	Val	Leu	Asn	Glu	Pro	Ala	His	Leu	Thr	Ala	Leu	Arg	Asp
	195					200						205			

Val Phe Asn Gly Trp Gly Ser Tyr Glu Arg Gly Glu Asp Phe Tyr Val
 210 215 220
 Asp Val Phe Arg Trp Ala Lys Ala Ala Asp Ser Thr Ala Arg Leu Tyr
 225 230 235 240
 Ile Asn Glu Tyr Asn Ile Ile Asn Asn Tyr Ala Asn Glu Gln Pro Thr
 245 250 255
 Arg Asn Tyr Tyr Lys Trp Ile Ile Ala Arg Leu Ile Ser Lys Gly Ala
 260 265 270
 Pro Ile Glu Gly Ile Gly Ile Gln Gly His Ile Ser Ala Pro Leu Pro
 275 280 285
 Ser Met Ser Glu Val Lys Ala Ala Leu Asp Glu Met Ala Val Phe Gly
 290 295 300
 Leu Pro Leu Ala Ile Thr Glu Tyr Asp Val Thr Gly Val Ser Glu Glu
 305 310 315 320
 Val Glu Ala Asn Phe Met Arg Asp Phe Leu Thr Met Val Phe Ser His
 325 330 335
 Pro Ala Val Glu Ser Phe Val Met Trp Gly Phe Trp Ser Gly Ala His
 340 345 350
 Trp Arg Asp Asn Ala Pro Leu Phe Arg Ala Asp Trp Ser Leu Lys Pro
 355 360 365
 Ser Gly Gln Val Phe Leu Asp Leu Val Phe Arg Arg Trp Trp Thr Asp
 370 375 380
 Thr Thr Gly Val Thr Gly Pro Asp Gly Ser Trp Ser Val Arg Gly Phe
 385 390 395 400
 Leu Gly Asp Tyr Val Glu Val Gln Val Gly Glu Val Ser Val Thr
 405 410 415
 Lys Ser Leu Arg Leu Glu Ser Pro Gln Asp Thr Thr Thr Leu Glu Val
 420 425 430
 Val Val Ser Ser Val Lys Val Gly Lys Pro Thr Glu Asp Val Leu
 435 440 445
 Arg Val Gln Gly Phe Gly Pro Asp Pro Phe Val Glu Gly Thr Ala Leu
 450 455 460
 Arg Tyr Trp Leu Gly Arg Pro Ala Asp Val Glu Leu Ala Val Tyr Asp
 465 470 475 480
 Val Leu Gly Arg Gln Val Tyr Ala Val Gln Lys His Arg Val Ala Gly
 485 490 495
 Trp His Thr Glu Trp Val Glu Ala Ser His Trp Pro Ala Gly Leu Tyr
 500 505 510
 Leu Tyr Arg Leu Gln Ala Gly Asp Leu Leu His Thr Gly Arg Met Val
 515 520 525
 Lys Ile Gln
 530

<210> 71
 <211> 1269
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 71
 atgatttcca tcgcccattc agtgaacggg gcggagacct atgagaacct ggcgcacgga 60
 actgccaaaa aaggcgaatg gactacgctg aaggggacat acaccgccgg cgcctatcag 120
 cgcaacgtgc tctatgtgga aacggtttct gaaggcacc ttgactttga gatccgtaat 180
 tttgtcctga cggctccgaa cggactaccg gagcccaagc cgaccgagcc tccgatgggtc 240
 atcgaggaag ccgagaacgt gccagctctc aaagagattt atgcagacaa attcgatttc 300
 ggctccgccg cgcccagat ggtattccgt gaccccaaat ggctcaacct gatgaaggaa 360
 cagttcagca ttctgacgcc ggaaaacgaa atgaaaccgg attccgttct ggatgtgggc 420
 gcgagcaaag cgctggtgaa ggaaaccggt gatgagaccg ccgtcgccgt tcatttcgac 480
 gctgccaaag cgctgctgaa ttttgccaag agcaacggga tcaaggttca cggccatgtg 540
 ctgatctggc acagccagac gccggaagct ttcttccatc agggatatga ttccaagaag 600
 ctttctgcta cacgggaagt gatgctgggc cgaatggaaa attacattaa ggggtgtttt 660
 gaatacctgg atgaaaatta tcccggcgct gttgtctcct gggacgtgct gaatgaggcg 720
 attgacgacg gaagcaactg gctgcggaac agcaactgga gaaagattgt cggcgaagac 780
 tatccgaacc gggcatatga atatgcgcgc aaatatgcgc cggaaggtac gctgctgtat 840
 tacaacgatt acaattacgtc gattcccggg aaactgaacg gcattgtgaa actgctgaac 900
 agtctgattc cggaaggaaa tatcgacggg tacggcttcc agatgcacca tggcgtcggc 960
 ttcccgctca ttgatatgat ccagactgca gtggaacgga ttgccgcgct gaatatccgc 1020

cttcgcgtca	gcgaactgga	tgctcacggtg	gacaacaaca	cggaagcgtc	cttcaacaaa	1080
caggcaaaagt	attatgccga	agtcatagaag	attctgattg	ctcacagcga	ccagtttgag	1140
gctgtgcagg	tctgggggct	gacagacctg	atgagctggc	gcggcagtc	gttcccgtg	1200
ctgtttgacg	gggcaggcaa	tccgaaaccg	gcgttctggtg	ccgtcgcgga	tccggattcc	1260
gtgaaataa						1269

<210> 72
 <211> 422
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 72
 Met Ile Ser Ile Ala His Ser Val Asn Gly Ala Glu Thr Tyr Glu Asn
 1 5 10 15
 Leu Ala His Gly Thr Ala Lys Lys Gly Glu Trp Thr Thr Leu Lys Gly
 20 25 30
 Thr Tyr Thr Ala Gly Ala Tyr Gln Arg Asn Val Leu Tyr Val Glu Thr
 35 40 45
 Val Ser Glu Gly Thr Leu Asp Phe Glu Ile Arg Asn Phe Val Leu Thr
 50 55 60
 Ala Pro Asn Gly Leu Pro Glu Pro Lys Pro Thr Thr Glu Pro Pro Met Val
 65 70 75 80
 Ile Glu Glu Ala Glu Asn Val Pro Ser Leu Lys Glu Ile Tyr Ala Asp
 85 90 95
 Lys Phe Asp Phe Gly Ser Ala Ala Pro Gln Met Val Phe Arg Asp Pro
 100 105 110
 Lys Trp Leu Asn Leu Met Lys Glu Gln Phe Ser Ile Leu Thr Pro Glu
 115 120 125
 Asn Glu Met Lys Pro Asp Ser Val Leu Asp Val Gly Ala Ser Lys Ala
 130 135 140
 Leu Val Lys Glu Thr Gly Asp Glu Thr Ala Val Ala Val His Phe Asp
 145 150 155 160
 Ala Ala Lys Ala Leu Leu Asn Phe Ala Lys Ser Asn Gly Ile Lys Val
 165 170 175
 His Gly His Val Leu Ile Trp His Ser Gln Thr Pro Glu Ala Phe Phe
 180 185 190
 His Gln Gly Tyr Asp Ser Lys Lys Pro Phe Val Thr Arg Glu Val Met
 195 200 205
 Leu Gly Arg Met Glu Asn Tyr Ile Lys Gly Val Phe Glu Tyr Leu Asp
 210 215 220
 Glu Asn Tyr Pro Gly Val Val Ser Trp Asp Val Leu Asn Glu Ala
 225 230 235 240
 Ile Asp Asp Gly Ser Asn Trp Leu Arg Asn Ser Asn Trp Arg Lys Ile
 245 250 255
 Val Gly Glu Asp Tyr Pro Asn Arg Ala Tyr Glu Tyr Ala Arg Lys Tyr
 260 265 270
 Ala Pro Glu Gly Thr Leu Leu Tyr Tyr Asn Asp Tyr Asn Thr Ser Ile
 275 280 285
 Pro Gly Lys Leu Asn Gly Ile Val Lys Leu Leu Asn Ser Leu Ile Pro
 290 295 300
 Glu Gly Asn Ile Asp Gly Tyr Gly Phe Gln Met His His Gly Val Gly
 305 310 315 320
 Phe Pro Ser Ile Asp Met Ile Gln Thr Ala Val Glu Arg Ile Ala Ala
 325 330 335
 Leu Asn Ile Arg Leu Arg Val Ser Glu Leu Asp Val Thr Val Asp Asn
 340 345 350
 Asn Thr Glu Ala Ser Phe Asn Lys Gln Ala Lys Tyr Tyr Ala Glu Val
 355 360 365
 Met Lys Ile Leu Ile Ala His Ser Asp Gln Phe Glu Ala Val Gln Val
 370 375 380
 Trp Gly Leu Thr Asp Leu Met Ser Trp Arg Gly Ser Gln Phe Pro Leu
 385 390 395 400
 Leu Phe Asp Gly Ala Gly Asn Pro Lys Pro Ala Phe Trp Ala Val Ala
 405 410 415
 Asp Pro Asp Ser Val Lys
 420

<210> 73
 <211> 4455
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 73
 atgcagaaaa tgagaagaaa attgaaaaga attatgttat tacttctggc agctatgttg 60
 ataattcccg caggctggat tacacaggct tcagcagcgg aaacaaacaa agatatacct 120
 gttctactgt accatcgaat tgttgataat cctactaatc aatggacgga taccagcggt 180
 gaaacgttta aacagactat gcaatatcta aatgatagcg gttacaacac ctgtcagccc 240
 gaacaatatg taaagatcat ggatggaacg gcaacggcgc ctgaaaaacc gattctatta 300
 acgtttgacg atgggtactcc agaatttatc accaatgctc ttccagtatt aaagcaatat 360
 aacatgaaag ctgttctgtt tattgtcagt gactggatag gcggcggcct cagcatgtca 420
 aaagaacagc tgcaaggttt ggctaatagaa ccattcttaa gcctcgaaaa tcatacgaaa 480
 acccatgacg gtactatttt gggaacaaat ggcggtgtac gtagtacgat aacgaaagaa 540
 caagctgagg accaaattat atcagcggaat acttatctta aaagtattac aggtaaagac 600
 ccagtcctaa ttgcataccc ttatggcagc tataatgata ttgcaaaact agtaaaccaa 660
 gaaaatggta ttaagtacgc atttaagtg ggataacccta atgaagataa ttatgctatg 720
 ggccgtcact atgtaacaaa tcaaagtgtg gctcaaattg cccaaatgat tggcggccct 780
 gtgccagAAC caactccaga accaggaaac cagacagaaa ccgtctatca agaaaccttt 840
 gccagtgata ttggtgtagc agttcaagcg ggtaaccac aagtaaccac cgttctgtgt 900
 atgggttttt caggcaatga cgatggaaaa gccatctctg tttagcggcag gacgaacaa 960
 tgggacggcg tcgatatccc attcaacaat gtcggtatgg aaaacggcaa aacttatacg 1020
 attacagtta ctggttatgt tgacgaaaat gcaactgttc cttctggcgc acaagcttta 1080
 ctgcagaatg tagacagcta taacggtttg tatgttgccg cagattatgc agcgggacag 1140
 gcttttactt taacgggtca gtataccgtg gatactagta aagatagagc cctacgtatc 1200
 caatcaaatg atgctgggaa aactgttccg ttttacattg gaaacatctt gattacaacg 1260
 aaaaaaacga ctgcccctga aacagataga gtggtatttc acgaaacatt tggaaatggt 1320
 gttggtgttg ctacacaagc gggaagtgcg aaattgactc ctgtttctga gcttgttttt 1380
 gaaggcaata gcgatggaaa agcaatttct gttaatggca gatccaataa ctgggacgga 1440
 gttgatatac ctttcagcag tgtcagcatg cagaacggca aagcctatac cattacagtc 1500
 actggtttttg ttatatagcag tgtgagtgtt cctgaagggtg cacaagcttt gcttcagaat 1560
 gtagacagct ataattggctt gtatgcagca gcagatgtta aggcagggtca aacatttact 1620
 ctaacgggtc aatataccgt tgatacagagc aaagatagag cactacgaat tcaatcaaat 1680
 gatgcaggga aaaccgttcc ctcttatatc ggagataattc tcattaccga gaaggcagcc 1740
 tctggtgggtg gcggggacga tggaaagacta cctgcccgaac catttacagc aatttaatttt 1800
 gaagaccaaa atctatgggtg ttctgagggg agagctggta ccgaaacact aacagtaacc 1860
 aatgaagcca atcatactga tggcgggttc tatgctttga aggttgaagg cagatcacia 1920
 gcttggcatg gaccagcatt acacgtagag aaatatgttg acaaggattc ggaatataaa 1980
 atttctgcct ggggtgaagc gatttcacca gcaactcac agcttcagct ttctacacag 2040
 gtccggcaatg gcggaaactg tagttacaat aatcttcaag gaaaaactat cagcactgaa 2100
 gatggctggg ttaaacttga gggaacgtac cgttatagca gtgtaggcga tgagttttta 2160
 accatttatg tagagagctc gaataatagc acagcctcct tttatatcga tgatattact 2220
 ttgaatcga ctggttccgg tccgattgaa gttgaggatt tgacaccgat aaaagatgtt 2280
 tatcaagacg atttcttaag tggaaacgct gtctcagctt ctgatcttga aggcaataga 2340
 cttaagcttc tcaacatgca tcacaatgtt gtcacagcag agaattgcaat gaagccagat 2400
 caagcgtata atgcggaaaa acaatttgac tttactgatg aaaatgcgct tgtcgacaag 2460
 gttttggatc agggattgca gctgcatggt cacgtgcttg tatggacca gcagacgcca 2520
 gaatgggttat ttacagtga aaacggtgac cctttgagcc gtgaggcagc actagcaaat 2580
 ttaaggaccc atgttaaaac agtcgtagaa aattacggta acaaggtaat ttcattgggac 2640
 gtggtaaacg aagcaatcat cgataaccgc ccgaacccaa cggattggaa ggcatacatt 2700
 cgtaaatctg gctggtacaa atcgattgga ccagacttcg tagaacaatc cttccttgct 2760
 gcaaaagagg tactgaatga aaaaggcttg aatatcaagc tatattacaa tgattacaat 2820
 gatgataatc agagcaagc cgaggccatt tatcagatgg tgaaagatat caatgaaaag 2880
 tatgtctaaag aacatgatgg ggaacttctc attgacggaa ttggaatgca agcgactac 2940
 aataaaaaaca ctaatcctga aaatgttaaa ctctccctag agaagtttat tacattgggt 3000
 gtagaagtca gtgtgactga acttgacatt accgctggaa ccaataatgt acttactgag 3060
 aaggaagcaa ttgcacaggg ttattttata gcacaattgt tcaagattta caaagaacac 3120
 gcagagcata tctcacgggt aactttctgg ggactaaatg atgcaacgag ctggagagct 3180
 gcacagagtc cattgttgtt tgataaagat ttgcaagcaa aaccagctta ctatgctgtt 3240
 atcgatccag acacattttac tgtagaaaat caacctgagg taagagaggc taatcaagga 3300
 agtgctgttt ccggcacacc agtgattgat ggaactgtag atgggtgttt gagcaatgca 3360
 acggaactgc cttccaaatg gtttggcagg gagcaaacgg ggtatccaag 3420
 gtccctctggg ataattgaaa cctgtatgtt ttaattcaag taagtgactc acagctcgac 3480
 aaatcgagtc caaatccatg ggaacaggat tccattgaag tctttgtaga tgagaataat 3540

gcaaagacat	cttccttcga	agatggtgat	ggacaatatc	gagtaaactt	tgacaatgaa	3600
acatccttta	accctgtcag	agttggagaa	ggtttcgaat	ctgcaaccaa	agcatcaggt	3660
aattggctata	ccgttgaagt	aaagattccg	ttcaaaaacca	ttacaccaga	taacaatacg	3720
aaaatcgggt	ttgatgttca	gattaatgac	ggtaaagatg	gtgctcgtca	aagtgtctgca	3780
acatggaacg	atttaactgg	tctgggatat	caggacactt	ctgtgttcgg	cgctcctgaca	3840
cttatgaaga	ctgacaccac	cgcgcctggt	acaaccgata	acggaccaga	agattggggt	3900
aataaagatg	taacgattgc	tttcagtgc	aatgataatg	acactgggtg	ggcggcaacc	3960
tattatagta	ttgataatgg	ggtcgtacaa	aacggtaatt	cagttactat	ttcgggaagag	4020
ggtgtccaca	ttctaacata	ttggagtgt	gacaaagctg	gtaatgtcga	gcaggttcat	4080
acaaaaacaa	ttaaactaga	taagaccgga	ccaatattag	atattaaact	cgacaaaaca	4140
acattatcac	cagttaatca	taagatggtc	ccaatatcgg	cggctattag	tgcattctgat	4200
gccgattcag	gaattcattc	agtagtggt	acatcaatta	ctagcaatga	atctatccaa	4260
cctgatgata	ttcagaatgc	caactataat	aaacctatta	cagggtactac	ggattccttt	4320
aaacttcgtg	cagaaagatt	agcaaacggt	aattggccgtg	tttacaccat	tacttatacg	4380
gccacagata	aagctggtaa	tgtgacaaca	aaaagtgttg	aagtttccgt	tccacgcgac	4440
aatttctaaaa	aataa					4455

<210> 74
 <211> 1484
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(21)

<400> 74

Met	Gln	Lys	Met	Arg	Arg	Lys	Leu	Lys	Arg	Ile	Met	Leu	Leu	Leu	Leu
1				5					10					15	
Ala	Ala	Met	Leu	Ile	Ile	Pro	Ser	Gly	Trp	Ile	Thr	Gln	Ala	Ser	Ala
		20						25					30		
Ala	Glu	Thr	Asn	Lys	Asp	Ile	Pro	Val	Leu	Leu	Tyr	His	Arg	Ile	Val
		35					40					45			
Asp	Asn	Pro	Thr	Asn	Gln	Trp	Thr	Asp	Thr	Ser	Val	Glu	Thr	Phe	Lys
	50					55					60				
Gln	Thr	Met	Gln	Tyr	Leu	Asn	Asp	Ser	Gly	Tyr	Asn	Thr	Leu	Ser	Ala
	65				70					75				80	
Glu	Gln	Tyr	Val	Lys	Ile	Met	Asp	Gly	Thr	Ala	Thr	Ala	Pro	Glu	Lys
			85					90					95		
Pro	Ile	Leu	Leu	Thr	Phe	Asp	Asp	Gly	Thr	Pro	Glu	Phe	Ile	Thr	Asn
		100						105				110			
Ala	Leu	Pro	Val	Leu	Lys	Gln	Tyr	Asn	Met	Lys	Ala	Val	Leu	Phe	Ile
		115					120					125			
Val	Ser	Asp	Trp	Ile	Gly	Gly	Gly	Phe	Ser	Met	Ser	Lys	Glu	Gln	Leu
	130				135						140				
Gln	Ser	Leu	Ala	Asn	Glu	Pro	Ser	Leu	Ser	Leu	Glu	Asn	His	Thr	Lys
	145			150					155					160	
Thr	His	Asp	Gly	Thr	Ile	Trp	Gly	Thr	Asn	Gly	Gly	Val	Arg	Ser	Thr
		165						170					175		
Ile	Thr	Lys	Glu	Gln	Ala	Glu	Asp	Gln	Ile	Ile	Ser	Ala	Asn	Thr	Tyr
		180						185				190			
Leu	Lys	Ser	Ile	Thr	Gly	Lys	Asp	Pro	Val	Leu	Met	Ala	Tyr	Pro	Tyr
		195					200					205			
Gly	Ser	Tyr	Asn	Asp	Ile	Ala	Lys	Leu	Val	Asn	Gln	Glu	Asn	Gly	Ile
	210				215						220				
Lys	Tyr	Ala	Phe	Lys	Val	Gly	Tyr	Pro	Asn	Glu	Asp	Asn	Tyr	Ala	Met
	225				230					235					240
Gly	Arg	His	Tyr	Val	Thr	Asn	Gln	Ser	Val	Ala	Gln	Ile	Ala	Gln	Met
			245						250					255	
Ile	Gly	Gly	Pro	Val	Pro	Glu	Pro	Thr	Pro	Glu	Pro	Gly	Asn	Gln	Thr
			260					265				270			
Glu	Thr	Val	Tyr	Gln	Glu	Thr	Phe	Ala	Ser	Asp	Ile	Gly	Val	Ala	Val
		275					280					285			
Gln	Ala	Gly	Asn	Pro	Gln	Val	Thr	His	Val	Ser	Gly	Met	Val	Phe	Ala
	290					295					300				
Gly	Asn	Asp	Asp	Gly	Lys	Ala	Ile	Ser	Val	Ser	Gly	Arg	Thr	Asn	Asn
	305				310					315					320

Trp Asp Gly Val Asp³²⁵ Ile Pro Phe Asn³³⁰ Val Gly Met Glu Asn³³⁵ Gly
 Lys Thr Tyr Thr³⁴⁰ Ile Thr Val Thr Gly³⁴⁵ Tyr Val Asp Glu Asn³⁵⁰ Ala Thr
 Val Pro Ser³⁵⁵ Gly Ala Gln Ala Leu³⁶⁰ Leu Gln Asn Val Asp³⁶⁵ Ser Tyr Asn
 Gly Leu³⁷⁰ Tyr Val Ala Ala Asp³⁷⁵ Tyr Ala Ala Gly Gln³⁸⁰ Ala Phe Thr Leu
 Thr³⁸⁵ Gly Gln Tyr Thr Val³⁹⁰ Asp Thr Ser Lys Asp³⁹⁵ Arg Ala Leu Arg Ile
 Gln Ser Asn Asp⁴⁰⁵ Ala Gly Lys Thr Val Pro⁴¹⁰ Phe Tyr Ile Gly Asn⁴¹⁵ Ile
 Leu Ile Thr Thr⁴²⁰ Lys Lys Thr Thr Ala⁴²⁵ Pro Glu Thr Asp Arg Val Val
 Phe His Glu⁴³⁵ Thr Phe Gly Asn Gly⁴⁴⁰ Val Gly Val Ala Thr⁴⁴⁵ Gln Ala Gly
 Ser Ala Lys⁴⁵⁰ Leu Thr Pro Val⁴⁵⁵ Ser Glu Leu Val Phe⁴⁶⁰ Glu Gly Asn Ser
 Asp⁴⁶⁵ Gly Lys Ala Ile Ser Val⁴⁷⁰ Asn Gly Arg Ser⁴⁷⁵ Asn Asn Trp Asp Gly
 Val Asp Ile Pro Phe⁴⁸⁵ Ser Ser Val Ser Met⁴⁹⁰ Gln Asn Gly Lys Ala Tyr
 Thr Ile Thr Val⁵⁰⁰ Thr Gly Phe Val Tyr⁵⁰⁵ Ser Ser Val Ser Val⁵¹⁰ Pro Glu
 Gly Ala Gln⁵¹⁵ Ala Leu Leu Gln Asn⁵²⁰ Val Asp Ser Tyr Asn⁵²⁵ Gly Leu Tyr
 Ala Ala Ala Asp Val Lys Ala Gly⁵³⁵ Gln Thr Phe Thr⁵⁴⁰ Leu Thr Gly Gln
 Tyr⁵⁴⁵ Thr Val Asp Thr Ser⁵⁵⁰ Lys Asp Arg Ala Leu⁵⁵⁵ Arg Ile Gln Ser Asn
 Asp Ala Gly Lys Thr⁵⁶⁵ Val Pro Phe Tyr Ile⁵⁷⁰ Gly Asp Ile Leu Ile Thr
 Glu Lys Ala Ala⁵⁸⁰ Ser Gly Gly Gly Gly⁵⁸⁵ Asp Asp Gly Arg Leu⁵⁹⁰ Pro Ala
 Glu Pro Phe⁵⁹⁵ Thr Ala Ile Asn Phe⁶⁰⁰ Glu Asp Gln Asn Met⁶⁰⁵ Gly Gly Phe
 Glu Gly Arg Ala Gly Thr Glu⁶¹⁵ Thr Leu Thr Val Thr⁶²⁰ Asn Glu Ala Asn
 His⁶²⁵ Thr Asp Gly Gly Ser⁶³⁰ Tyr Ala Leu Lys Val⁶³⁵ Glu Gly Arg Ser Gln
 Ala Trp His Gly Pro⁶⁴⁵ Ala Leu His Val Glu⁶⁵⁰ Lys Tyr Val Asp Lys Asp
 Ser Glu Tyr Lys⁶⁶⁰ Ile Ser Ala Trp Val⁶⁶⁵ Lys Leu Ile Ser Pro⁶⁷⁰ Ala Thr
 Ser Gln Leu Gln Leu Ser Thr Gln⁶⁸⁰ Val Gly Asn Gly Gly⁶⁸⁵ Thr Ala Ser
 Tyr Asn⁶⁹⁰ Asn Leu Gln Gly Lys⁶⁹⁵ Thr Ile Ser Thr Glu⁷⁰⁰ Asp Gly Trp Val
 Lys⁷⁰⁵ Leu Glu Gly Thr Tyr⁷¹⁰ Arg Tyr Ser Ser Val⁷¹⁵ Gly Asp Glu Phe Leu
 Thr Ile Tyr Val Glu⁷²⁵ Ser Ser Asn Asn Ser⁷³⁰ Thr Ala Ser Phe Tyr Ile
 Asp Asp Ile Thr⁷⁴⁰ Phe Glu Ser Thr Gly⁷⁴⁵ Ser Gly Pro Ile Glu Val Glu
 Asp Leu Thr⁷⁵⁵ Pro Ile Lys Asp Val⁷⁶⁰ Tyr Gln Asp Asp Phe⁷⁶⁵ Leu Ile Gly
 Asn Ala Val Ser Ala Ser Asp⁷⁷⁵ Leu Glu Gly Asn Arg⁷⁸⁰ Leu Lys Leu Leu
 Asn⁷⁸⁵ Met His His Asn Val Val Thr Ala Glu Asn Ala Met Lys Pro Asp
 Gln Ala Tyr Asn Ala⁸⁰⁵ Glu Lys Gln Phe Asp⁸¹⁰ Phe Thr Asp Glu Asn Ala
 Leu Val Asp Lys⁸²⁰ Val Leu Asp Gln Gly⁸²⁵ Leu Gln Leu His Gly His Val
 Leu Val Trp⁸³⁵ His Gln Gln Thr Pro⁸⁴⁰ Glu Trp Leu Phe Thr Ala Glu Asn
 Gly Ala⁸⁵⁰ Pro Leu Ser Arg Glu⁸⁵⁵ Ala Ala Leu Ala Asn⁸⁶⁰ Leu Arg Thr His
 Val Lys Thr Val Val Glu Asn Tyr Gly Asn Lys Val Ile Ser Trp Asp

865 Val Val Asn Glu Ala 870 Ile Ile Asp Asn Pro 875 Pro Asn Pro Thr Asp 880 Trp
 Lys Ala Ser Leu Arg Lys Ser Gly Trp Tyr Lys Ser Ile Gly Pro Asp
 Phe Val Glu 900 Ser Phe Leu Ala Lys Glu Val Leu Asn Glu Lys
 Gly Leu Asn Ile Lys Leu Tyr Tyr Asn Asp Tyr Asn Asp Asp Asn Gln
 Ser Lys Ala Glu Ala Ile Tyr Gln Met Val Lys Asp Ile Asn Glu Lys
 Tyr Ala Lys Glu His 950 Asp Gly Asp Leu Leu Ile Asp Gly Ile Gly Met
 Gln Ala His Tyr Asn Lys Asn Thr Asn Pro Glu Asn Val Lys Leu Ser
 Leu Glu Lys Phe Ile Thr Leu Gly Val Glu Val Ser Val Thr Glu Leu
 Asp Ile Thr Ala Gly Thr Asn Asn Val Leu Thr Glu Lys Glu Ala Ile
 Ala Gln Gly Tyr Leu Tyr Ala Gln Leu Phe Lys Ile Tyr Lys Glu His
 Ala Glu His Ile Ser Arg Val Thr Phe Trp Gly Leu Asn Asp Ala Thr
 Ser Trp Arg Ala Ala Gln Ser Pro Leu Leu Phe Asp Lys Asp Leu Gln
 Ala Lys Pro Ala Tyr Tyr Ala Val Ile Asp Pro Asp Thr Phe Thr Val
 Glu Asn Gln Pro Glu Val Arg Glu Ala Asn Gln Gly Ser Ala Val Ser
 Gly Thr Pro Val Ile Asp Gly Thr Val Asp Gly Val Trp Ser Asn Ala
 Thr Glu Leu Pro Ile Asn Arg Phe Gln Met Ala Trp Gln Gly Ala Asn
 Gly Val Ser Lys Val Leu Trp Asp Asn Glu Asn Leu Tyr Val Leu Ile
 Gln Val Ser Asp Ser Gln Leu Asp Lys Ser Ser Pro Asn Pro Trp Glu
 Gln Asp Ser Ile Glu Val Phe Val Asp Glu Asn Asn Ala Lys Thr Ser
 Ser Phe Glu Asp Gly Asp Gly Gln Tyr Arg Val Asn Phe Asp Asn Glu
 Thr Ser Phe Asn Pro Val Arg Val Gly Glu Gly Phe Glu Ser Ala Thr
 Lys Ala Ser Gly Asn Gly Tyr Thr Val Glu Val Lys Ile Pro Phe Lys
 Thr Ile Thr Pro Asp Asn Asn Thr Lys Ile Gly Phe Asp Val Gln Ile
 Asn Asp Gly Lys Asp Gly Ala Arg Gln Ser Ala Ala Thr Trp Asn Asp
 Leu Thr Gly Leu Gly Tyr Gln Asp Thr Ser Val Phe Gly Val Leu Thr
 Leu Met Lys Thr Asp Thr Thr Ala Pro Val Thr Thr Asp Asn Gly Pro
 Glu Asp Trp Val Asn Lys Asp Val Thr Ile Ala Phe Ser Ala Asn Asp
 Asn Asp Thr Gly Val Ala Ala Thr Tyr Tyr Ser Ile Asp Asn Gly Val
 Val Gln Asn Gly Asn Ser Val Thr Ile Ser Glu Glu Gly Val His Ile
 Leu Thr Tyr Trp Ser Val Asp Lys Ala Gly Asn Val Glu Gln Val His
 Thr Lys Thr Ile Lys Leu Asp Lys Thr Gly Pro Ile Leu Asp Ile Lys
 Leu Asp Lys Thr Thr Leu Ser Pro Val Asn His Lys Met Val Pro Ile
 Ser Ala Ala Ile Ser Ala Ser Asp Ala Asp Ser Gly Ile His Ser Val
 Val Leu Thr Ser Ile Thr Ser Asn Glu Ser Ile Gln Pro Asp Asp Ile
 1410 1415 1420

Gln Asn Ala Asn Tyr Asn Lys Pro Ile Thr Gly Thr Thr Asp Ser Phe
 1425 1430 1435 1440
 Lys Leu Arg Ala Glu Arg Leu Ala Asn Gly Asn Gly Arg Val Tyr Thr
 1445 1450 1455
 Ile Thr Tyr Thr Ala Thr Asp Lys Ala Gly Asn Val Thr Thr Lys Ser
 1460 1465 1470
 Val Glu Val Ser Val Pro Arg Asp Asn Ser Lys Lys
 1475 1480

<210> 75
 <211> 1122
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 75
 atgaaaaagc atattgtact cttcgcatct ctttcctgta ttttgctggc ggccccgatcg 60
 tcggcgctcg agcgatttct caaggacgtc ttttcggatt ccttcaaggc cggcgtagcc 120
 ctcaatgccg atcagattac gggggcggac tcggccagcc tcgacttgct cttggctcac 180
 ttcgatttct ttgtcgctga aaatgcgatg aagtgggggt cgctcaatcc tgagccgggg 240
 gtttacgatt tccgggtggc tgacgccctg gtcgatttgg cggagcggga aggtttgttt 300
 ttggttggcc acacactgct ctggcatcag cagacgccgg actgggtttt tctggacgag 360
 aagggcgaga ccgccacgcg ggagctggtg ctgcctcgac tggagacgca catccgcacc 420
 gtggtcggcc gctaccaggg ccgggtgcag ggctgggatg tggtaacga agccttgaac 480
 gaagacggtt cgttgcggga gtcgaaatgg ttgcagatca tcggcccggga ctacatcgaa 540
 ctggcggtcc gcatggcgaa ggagggcgat cccgacgccg agctttatta caatgactac 600
 aatgtgtcca agcccggcaa gcgaggtgga gtggtgcgcc tgcttggaga gctgcaggcg 660
 aaaggagtta aggtcgatgc ggtcggcatc cagggccact acagtctcgg gcaccctgag 720
 ctcgaccagc tcgaggccag catttctgcg ataacggagg ctggggctcc gatcatgata 780
 accgagctcg atgtgtcggg cttgcccttt cccgacgcgg agcaaattggg ggcggacgtg 840
 tcgctcagct ttgagatgca ggaccacctc aatccctatg ccgatggctt gcccgaggcg 900
 gtttcgcagc agctagctga acgttacgcg gccatttttg aagtgttttt gcgccaccag 960
 agccacatcg accgcgtgac gttttgggga gtgcacgac gggtcagctg gtggaactat 1020
 tggccgatcg cgggcaggac cgactatccc ttgctgtttg atcgggagct caagcggaaa 1080
 gcggccttcg aggcggtggt cgatttagcg gagggccgct ga 1122

<210> 76
 <211> 373
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(22)

<400> 76
 Met Lys Lys His Ile Val Leu Phe Ala Phe Leu Ser Val Ile Leu Leu
 1 5 10 15
 Ala Ala Arg Ser Ala Ser Glu Arg Phe Leu Lys Asp Val Phe Ser
 20 25 30
 Asp Ser Phe Lys Val Gly Val Ala Leu Asn Ala Asp Gln Ile Thr Gly
 35 40 45
 Ala Asp Ser Ala Ser Leu Asp Leu Ser Leu Ala His Phe Asp Ser Leu
 50 55 60
 Val Ala Glu Asn Ala Met Lys Trp Gly Ser Leu Asn Pro Glu Pro Gly
 65 70 75 80
 Val Tyr Asp Phe Arg Val Ala Asp Ala Leu Val Asp Leu Ala Glu Arg
 85 90 95
 Glu Gly Leu Phe Leu Val Gly His Thr Leu Leu Trp His Gln Gln Thr
 100 105 110
 Pro Asp Trp Val Phe Leu Asp Glu Lys Gly Glu Thr Ala Thr Arg Glu
 115 120 125
 Leu Val Leu Ala Arg Leu Glu Thr His Ile Arg Thr Val Val Gly Arg
 130 135 140
 Tyr Gln Gly Arg Val Gln Gly Trp Asp Val Val Asn Glu Ala Leu Asn

145 150 155 160
 Glu Asp Gly Ser Leu Arg Glu Ser Lys Trp Leu Gln Ile Ile Gly Pro
 165 170 175
 Asp Tyr Ile Glu Leu Ala Phe Arg Met Ala Lys Glu Ala Asp Pro Asp
 180 185 190
 Ala Glu Leu Tyr Tyr Asn Asp Tyr Asn Val Ser Lys Pro Gly Lys Arg
 195 200 205
 Gly Gly Val Val Arg Leu Leu Gly Glu Leu Gln Ala Lys Gly Val Lys
 210 215 220
 Val Asp Ala Val Gly Ile Gln Gly His Tyr Ser Leu Gly His Pro Glu
 225 230 235 240
 Leu Asp Gln Leu Glu Ala Ser Ile Ser Ala Ile Thr Glu Ala Gly Ala
 245 250 255
 Pro Ile Met Ile Thr Glu Leu Asp Val Ser Val Leu Pro Phe Pro Asp
 260 265 270
 Ala Glu Gln Met Gly Ala Asp Val Ser Leu Ser Phe Glu Met Gln Asp
 275 280 285
 His Leu Asn Pro Tyr Ala Asp Gly Leu Pro Glu Ala Val Ser Gln Gln
 290 295 300
 Leu Ala Glu Arg Tyr Ala Ala Ile Phe Glu Val Phe Leu Arg His Gln
 305 310 315 320
 Ser His Ile Asp Arg Val Thr Phe Trp Gly Val His Asp Gly Val Ser
 325 330 335
 Trp Trp Asn Tyr Trp Pro Ile Ala Gly Arg Thr Asp Tyr Pro Leu Leu
 340 345 350
 Phe Asp Arg Glu Leu Lys Arg Lys Ala Ala Phe Glu Ala Val Val Asp
 355 360 365
 Leu Ala Glu Gly Arg
 370

<210> 77
 <211> 1248
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 77
 atgctaaaag ttttacgtaa acctatcggt tctggattag ctctagcctt attattacct 60
 ataggatcga cagttagtgc cgaaacaaat atttcaaata aaccagggtat tagcgggtta 120
 acagcaccac aattggacca acgatataaa gattctttca ccatagggtgc agcgggtgag 180
 ccaaatacaat tattagatgc aaaagactca caaatgttaa agcgccattt taatagcatt 240
 gtagcagaaa atgtcatgaa gcctagcagt ttacagccag tagaagggca gtttaactgg 300
 gaaccggcag ataaacttgt taagtttgcg aaagaaaatg gaatggacat gcgcggccat 360
 acgcttgtct ggcatagcca agtaccagat tggttcttca aagatgcaaa tggaaattca 420
 atggttggtt ggcagaatgg aaagcaagtg gttgcagatc cgtcaaatct tgaggctaac 480
 aaaaagcttt tattaagccg tttagaaaca catgttaata cagtcgtttc tcgttataaa 540
 aatgatatta aattttggga cgttgtaaat gaagtaatcg acgaatgggg cggacatcct 600
 gaaggtttac gtcaatctcc atggttccta attaccggaa cggactatat taaagtcgct 660
 tttgagacag caagacaata tgctgctcca gacgctaagc ttatatcaa tgattacaat 720
 acagaagtaa caccaaaaag aacgtactta tacaacctag taaaaagttt aaaacagcaa 780
 ggtgtttcaa ttgatggtgt tgggcatcag tctcacattc aaatcggctg gccgtctgaa 840
 aaagaaattg aagacacaat taacatgttt gctgaactgg ggttagacaa ccaaattact 900
 gagcttgatg taagcatgta tggctggcca gtaaggcggt atcctaccta tgattctatt 960
 ccagcacaga aatttataga tcaagcagac cgatatgatc gtttatttaa attatatgag 1020
 aaattaggcg ataaaatcag caatgtgaca ttctggggaa ttgctgataa ccatacatgg 1080
 ttaaattgacc gtgcagatgt ttactatgat gcagatggaa acgttgtaac attggcaaat 1140
 gcaccatatg ctaaaatgga agctagatca ggtaaagatg caccatttgt atttgatcca 1200
 gaatacaatg taaaaccagc ctattgggagc attatcgacc acaaataa 1248

<210> 78
 <211> 415
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
<222> (1)...(27)

<400> 78
Met Leu Lys Val Leu Arg Lys Pro Ile Val Ser Gly Leu Ala Leu Ala
1 5 10 15
Leu Leu Leu Pro Ile Gly Ser Thr Val Ser Ala Glu Thr Asn Ile Ser
20 25 30
Asn Lys Pro Gly Ile Ser Gly Leu Thr Ala Pro Gln Leu Asp Gln Arg
35 40 45
Tyr Lys Asp Ser Phe Thr Ile Gly Ala Ala Val Glu Pro Asn Gln Leu
50 55 60
Leu Asp Ala Lys Asp Ser Gln Met Leu Lys Arg His Phe Asn Ser Ile
65 70 75 80
Val Ala Glu Asn Val Met Lys Pro Ser Ser Leu Gln Pro Val Glu Gly
85 90 95
Gln Phe Asn Trp Glu Pro Ala Asp Lys Leu Val Lys Phe Ala Lys Glu
100 105 110
Asn Gly Met Asp Met Arg Gly His Thr Leu Val Trp His Ser Gln Val
115 120 125
Pro Asp Trp Phe Phe Lys Asp Ala Asn Gly Asn Ser Met Val Val Trp
130 135 140
Gln Asn Gly Lys Gln Val Val Ala Asp Pro Ser Asn Leu Glu Ala Asn
145 150 155 160
Lys Lys Leu Leu Ser Arg Leu Glu Thr His Val Asn Thr Val Val
165 170 175
Ser Arg Tyr Lys Asn Asp Ile Lys Phe Trp Asp Val Val Asn Glu Val
180 185 190
Ile Asp Glu Trp Gly Gly His Pro Glu Gly Leu Arg Gln Ser Pro Trp
195 200 205
Phe Leu Ile Thr Gly Thr Asp Tyr Ile Lys Val Ala Phe Glu Thr Ala
210 215 220
Arg Gln Tyr Ala Ala Pro Asp Ala Lys Leu Tyr Ile Asn Asp Tyr Asn
225 230 235 240
Thr Glu Val Thr Pro Lys Arg Thr Tyr Leu Tyr Asn Leu Val Lys Ser
245 250 255
Leu Lys Gln Gln Gly Val Pro Ile Asp Gly Val Gly His Gln Ser His
260 265 270
Ile Gln Ile Gly Trp Pro Ser Glu Lys Glu Ile Glu Asp Thr Ile Asn
275 280 285
Met Phe Ala Glu Leu Gly Leu Asp Asn Gln Ile Thr Glu Leu Asp Val
290 295 300
Ser Met Tyr Gly Trp Pro Val Arg Ala Tyr Pro Thr Tyr Asp Ser Ile
305 310 315 320
Pro Ala Gln Lys Phe Ile Asp Gln Ala Asp Arg Tyr Asp Arg Leu Phe
325 330 335
Lys Leu Tyr Glu Lys Leu Gly Asp Lys Ile Ser Asn Val Thr Phe Trp
340 345 350
Gly Ile Ala Asp Asn His Thr Trp Leu Asn Asp Arg Ala Asp Val Tyr
355 360 365
Tyr Asp Ala Asp Gly Asn Val Val Thr Leu Ala Asn Ala Pro Tyr Ala
370 375 380
Lys Met Glu Ala Arg Ser Gly Lys Asp Ala Pro Phe Val Phe Asp Pro
385 390 395 400
Glu Tyr Asn Val Lys Pro Ala Tyr Trp Ala Ile Ile Asp His Lys
405 410 415

<210> 79
<211> 1293
<212> DNA
<213> Unknown

<220>
<223> obtained from an environmental sample

<400> 79
atgattggtc tggatttgat ttctgggtggt cgtcgcaagg cctgtctggc tgcctgtctg 60
gcgcttgccg cgctgtcatt gccgggtatcg gctcaaatgg ctgcggggaa ggaaaagttc 120
gtgggtaacg tgatcgctgg ttatgtgccc ggtgattacg gcaatctctg gaatcaggtg 180

acgccggaga	attccaccaa	gtggggagcg	gttgagtcta	cgcgtaatgt	catgaactgg	240
acgcaggctg	atctggccta	caactacgcc	aagtccaagg	gcttcaagtt	caagatgcac	300
acgctggtat	ggggctcgca	agagccggcc	tgggtcaaga	atctggatgc	gacttcccag	360
cgtgtcgagg	tcgaacagtg	gatgcgtctg	agctgcgaac	gctaccccga	ttcctgggct	420
atcgatgtgg	tgaatgaacc	cctgcatgcc	gtgccctcgt	acaagaacgc	actgggtggc	480
gatggtgcca	ccggctggga	ttgggtcatc	acctcgttcc	gtctggcgcg	tcagtactgt	540
ccgcgcgcca	agctgctgct	caatgagtac	gccaccgagc	tggatgccag	caagcgcgcc	600
aagatcaaga	ccattgcctc	gctgctcaag	agtcgcggtc	tgattgatgg	tgttggcctg	660
caggccatt	tcttcacgct	ggattacatg	aatgccagcc	agatgaaggc	ggcactggat	720
gattacgcca	cgctgggtgt	ggatatctac	atttccgagc	tggatctgaa	gggcagtgcc	780
aataccgacg	ccagccagaa	ggcgaagtac	gaagagctgt	tcccgggtgat	gtggaatcac	840
gccagcgtag	agggcatcac	cctgtggggc	tacaagggtg	gtgaaacctg	gtcgagcggc	900
accggcctgc	tgaatgcgaa	cggtagcgag	cgtccggccc	tgacctggct	gaaaagctat	960
atgagcagcc	gtcctgcagc	atcgagcagc	agttcttcga	gtgtttcatc	cagcaaattcc	1020
agttcgtctt	cttctagcca	gtccagtgcc	tccagcagtg	caggcagtg	gccggtcttg	1080
tccggcacca	gtgattacc	gagcgggttc	agcaagtgtg	ccgatctggg	cggcacttgc	1140
agcgtgtctt	ccggcaccgg	ctgggcggcc	ttcgggcgca	agggtaagtg	ggttgccaaa	1200
tacgtcggtg	tgggcaagag	cattccctgc	acggtggcgg	cgtttggtcg	tgacccgggg	1260
ggcaatccca	acaagtgttc	cttccagagg	taa			1293

<210> 80
 <211> 430
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(36)

<400> 80
 Met Ile Gly Leu Asp Leu Ile Ser Gly Gly Arg Arg Lys Ala Cys Leu
 1 5 10 15
 Ala Ala Cys Leu Ala Leu Ala Ala Leu Ser Leu Pro Val Ser Ala Gln
 20 25 30
 Met Ala Ala Gly Lys Glu Lys Phe Val Gly Asn Val Ile Ala Gly Tyr
 35 40 45
 Val Pro Gly Asp Tyr Gly Asn Leu Trp Asn Gln Val Thr Pro Glu Asn
 50 55 60
 Ser Thr Lys Trp Gly Ala Val Glu Ser Thr Arg Asn Val Met Asn Trp
 65 70 75 80
 Thr Gln Ala Asp Leu Ala Tyr Asn Tyr Ala Lys Ser Lys Gly Phe Lys
 85 90 95
 Phe Lys Met His Thr Leu Val Trp Gly Ser Gln Glu Pro Ala Trp Val
 100 105 110
 Lys Asn Leu Asp Ala Thr Ser Gln Arg Val Glu Val Glu Gln Trp Met
 115 120 125
 Arg Leu Ser Cys Glu Arg Tyr Pro Asp Ser Trp Ala Ile Asp Val Val
 130 135 140
 Asn Glu Pro Leu His Ala Val Pro Ser Tyr Lys Asn Ala Leu Gly Gly
 145 150 155 160
 Asp Gly Ala Thr Gly Trp Asp Trp Val Ile Thr Ser Phe Arg Leu Ala
 165 170 175
 Arg Gln Tyr Cys Pro Arg Ala Lys Leu Leu Leu Asn Glu Tyr Ala Thr
 180 185 190
 Glu Leu Asp Ala Ser Lys Arg Ala Lys Ile Lys Thr Ile Ala Ser Leu
 195 200 205
 Leu Lys Ser Arg Gly Leu Ile Asp Gly Val Gly Leu Gln Ala His Phe
 210 215 220
 Phe Thr Leu Asp Tyr Met Asn Ala Ser Gln Met Lys Ala Ala Leu Asp
 225 230 235 240
 Asp Tyr Ala Thr Leu Gly Val Asp Ile Tyr Ile Ser Glu Leu Asp Leu
 245 250 255
 Lys Gly Ser Ala Asn Thr Asp Ala Ser Gln Lys Ala Lys Tyr Glu Glu
 260 265 270
 Leu Phe Pro Val Met Trp Asn His Ala Ser Val Lys Gly Ile Thr Leu
 275 280 285
 Trp Gly Tyr Lys Val Gly Glu Thr Trp Ser Ser Gly Thr Gly Leu Leu

290 295 300
 Asn Ala Asn Gly Ser Glu Arg Pro Ala Leu Thr Trp Leu Lys Ser Tyr
 305 310 315 320
 Met Ser Ser Arg Pro Ala Ala Ser Ser Ser Ser Ser Ser Val Ser
 325 330 335
 Ser Ser Lys Ser Ser Ser Ser Ser Ser Gln Ser Ser Ala Ser Ser
 340 345 350
 Ser Ala Gly Ser Ala Pro Val Leu Ser Gly Thr Ser Asp Tyr Pro Ser
 355 360 365
 Gly Phe Ser Lys Cys Ala Asp Leu Gly Gly Thr Cys Ser Val Ser Ser
 370 375 380
 Gly Thr Gly Trp Ala Ala Phe Gly Arg Lys Gly Lys Trp Val Ala Lys
 385 390 395 400
 Tyr Val Gly Val Gly Lys Ser Ile Pro Cys Thr Val Ala Ala Phe Gly
 405 410 415
 Arg Asp Pro Gly Asn Pro Asn Lys Cys Ser Phe Gln Arg
 420 425 430

<210> 81
 <211> 1017
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 81
 ttgaccacga gagctattcg cacggaggca gcgctgaagg agatgtttgc ggaggacttt 60
 cagatcggag ccgctgttaa tccgatgact atacggacac aggaggagct gcttgcttat 120
 cacttcaaca gtattacggc agagaatgaa atgaagtgtg ccagtctgca gccggaggag 180
 ggggcttatg cttttgacga ggcggatcga ttggcggcct tcgcccggaa gcatggcatg 240
 gcgatgcggg gacacacttt agtgtggcat aaccagtcca caggctggct gttcgaagac 300
 aagcagggaa atcctgtaga taaggcaact ctgctggaga ggctgaaatc gcacatccat 360
 acggtagtag gacgttataa aaacgatatt tatgcttggg atgtggtaaa cgaggttata 420
 gaggacgagg gagacggcct gctgcgccgg tcgaaatggc tggatattgc cggaccggaa 480
 ttcattgccc gggcggttca gtatgctcat gaggctgacc ctaatgcgct gctcttctat 540
 aatgactaca acgagtccaa tccggcgaag cgagacaaga tccatgctct ggtgaagtcg 600
 ctgctggagc aaggcggtgcc tattcatggc attggactgc aggcgcattg gaatttgat 660
 ggtccttctc tcggcgagat ccgagcggca ctggagaagt atgcttctct tggcctgcag 720
 ctgcagctta cggagctgga tatgtcgtg tttcgttttg acgacaagcg tacggatata 780
 accgagcctc cggcggaatt gcttgagctg caggctgagc ggtatgagga aattttcaag 840
 ctgctgaggg aataccggga tgtaatcact tccgtgacct tctggggggc tgcggatgat 900
 tatacgtggc tgaacgattt tcccgtccgg gggcggaata attggccttt cctgttcgat 960
 gagcagcatc accccaaact ggcatttcat cgggtcgtg cactttccc cagtgaa 1017

<210> 82
 <211> 338
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 82
 Leu Thr Thr Arg Ala Ile Arg Thr Glu Ala Ala Leu Lys Glu Met Phe
 1 5 10 15
 Ala Glu Asp Phe Gln Ile Gly Ala Ala Val Asn Pro Met Thr Ile Arg
 20 25 30
 Thr Gln Glu Glu Leu Leu Ala Tyr His Phe Asn Ser Ile Thr Ala Glu
 35 40 45
 Asn Glu Met Lys Phe Ala Ser Leu Gln Pro Glu Glu Gly Ala Tyr Ala
 50 55 60
 Phe Asp Glu Ala Asp Arg Leu Ala Ala Phe Ala Arg Lys His Gly Met
 65 70 75 80
 Ala Met Arg Gly His Thr Leu Val Trp His Asn Gln Ser Thr Gly Trp
 85 90 95
 Leu Phe Glu Asp Lys Gln Gly Asn Pro Val Asp Lys Ala Thr Leu Leu
 100 105 110
 Glu Arg Leu Lys Ser His Ile His Thr Val Val Gly Arg Tyr Lys Asn

115 120 125
 Asp Ile Tyr Ala Trp Asp Val Asn Glu Val Ile Glu Asp Glu Gly
 130 135 140
 Asp Gly Leu Leu Arg Arg Ser Lys Trp Leu Asp Ile Ala Gly Pro Glu
 145 150 155 160
 Phe Ile Ala Arg Ala Phe Glu Tyr Ala His Glu Ala Asp Pro Asn Ala
 165 170 175
 Leu Leu Phe Tyr Asn Asp Tyr Asn Glu Ser Asn Pro Ala Lys Arg Asp
 180 185 190
 Lys Ile His Ala Leu Val Lys Ser Leu Leu Glu Gln Gly Val Pro Ile
 195 200 205
 His Gly Ile Gly Leu Gln Ala His Trp Asn Leu Tyr Gly Pro Ser Leu
 210 215 220
 Gly Glu Ile Arg Ala Ala Leu Glu Lys Tyr Ala Ser Leu Gly Leu Gln
 225 230 235 240
 Leu Gln Leu Thr Glu Leu Asp Met Ser Leu Phe Arg Phe Asp Asp Lys
 245 250 255
 Arg Thr Asp Ile Thr Glu Pro Pro Ala Glu Leu Leu Glu Leu Gln Ala
 260 265 270
 Glu Arg Tyr Glu Glu Ile Phe Lys Leu Leu Arg Glu Tyr Arg Asp Val
 275 280 285
 Ile Thr Ser Val Thr Phe Trp Gly Ala Ala Asp Asp Tyr Thr Trp Leu
 290 295 300
 Asn Asp Phe Pro Val Arg Gly Arg Lys Asn Trp Pro Phe Leu Phe Asp
 305 310 315 320
 Glu Gln His His Pro Lys Leu Ala Phe His Arg Val Ala Ala Leu Ser
 325 330 335
 Arg Gln

<210> 83
 <211> 3024
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 83
 atgaaaacca aaaggtctat attcaggttg tctatcctgg ttgtcctggc tgtgctgctg 60
 ttcagcgcaa tcaccttgac agccagcgcc gccgacacgc tcggcgcggc ggcggcccag 120
 tcgggcccgtt acttcggcac ggcgatagct gccggcaagc tcggcgactc gacctacacg 180
 accattgccca accgtgagtt caacatgatc acggctgaga atgagatgaa gatcgacgcc 240
 accgagccga accagaacca attcaacttc accaacgccg accggatctt caactgggcg 300
 gtgcagaatg ggaagcaggt gcgcgggcac acgctggcat ggcactcgca gcagccgggg 360
 tggatgagca gcatgagcgg caccgcgctg cgcaatgcga tgatcaacca catcaatggg 420
 gtgatggccc actacaaggg caggatctac gcctgggatg ttgtgaacga ggctttcaac 480
 gaggacggca gccgccgcaa ctggaacctg cagcagaccg gcaacgactg gatcgaggtg 540
 gccttcgcga cagcccgcac cgccgaccg gccgccaagc tgtgctacaa cgactacaac 600
 atcgaagcct ggagctatgc caagacgcag ggcgtttacc ggatgggtcca ggacttcaag 660
 tcccgcggcg tgccgatcga ctgtgtcggg ttccagagcc acttcaacag cggcacttcc 720
 tacgtcaaca gcaacttccg gacgacgctg caaagcttcg ccgcgctggg cgtggacgtg 780
 cagatcaccg agctggatgt cgagaatgcc gactcgcggc tcgattgggt gagaggcatc 840
 gtcaatgact gcctggcggt cccgcgctgc aacggcatca cgggtgaggg cgtgcgcgac 900
 agcgattcgt ggcgctcttc gcagaaccg ctgctgttca actccagcgg ttgtaagaag 960
 gcttcgtaca ccgccgtcct cgacgccttc aacgctgccc cgaccgtcac acctccggtg 1020
 acgacacctc cgggtgacgac accgccagtg accacgcctc ctcccgccac tgtgtcgatt 1080
 aacgcgggcg gctcggcgag cggcagcttc accggcgacc agtacttcag cggtggcagc 1140
 acctacacca acaccgccac catcgacatg agtcagatca ccagcaacc accgcccgcg 1200
 gcggtcttca acagcgagcg ttacggggcg atgacctaca ccatcccaa ccgctcgggt 1260
 gctcagacgg tgacgctgta ctttgccgag acctacctca ccgcggcagg gcagcggctg 1320
 ttcaacgtgt cgattaatgg cgagcggcg ctgtccaact tcgacatcta tgcctcgga 1380
 ggtggcgcta accgggcat cgcccggacg ttacgacca cggctaactc aagtggccag 1440
 gtgggtgatcc agttcacggc ggttacggag aaccccaaga tcaacgctat cacggtaaca 1500
 gcgggtggca cgcctccacc gacaacgcct ccgcccacca cgccgccacc gaccacctct 1560
 ccggtgacga cacccccagt gacgacaccc ccagtgcga caccgcccc cggcagcgtg 1620
 tcgatcaacg cgggcggctc ggccaccggc agcttcacgg gcgaccagta ctttagcgg 1680
 ggagcacct acaccaacac cgccaccatc gacatgagcc agatcaccag caaccacca 1740
 ccggcggcgg tgttcaacag cgagcgctac ggggcgatga cctacaccat ccccgccgc 1800

tcgggggctc	agacgggtcac	gctttacttt	gccgaaacgt	atgtcactgc	ggcagggcag	1860
cgctctttta	acgtgtctgt	aaacgggcgcg	gcagcgctgt	ccaacttcga	catctatgcc	1920
agcgccggcg	gccagaaccg	ggccatcgct	cgctccttca	acaccacggc	caactcaagc	1980
ggccaggtgg	tgatccagtt	cacggcggtc	accgagaacc	ccaagatcaa	cgccatcact	2040
gtggcgggcg	ggatcgggga	cttccaaacc	ctgaccgtca	cgaagtccgg	cacggggacg	2100
gtcacctcca	acccggctgg	tatcaactgc	ggctcgacct	gcaacgccag	cttcgctacc	2160
ggcaccagcg	tgaccctgac	cgcttcgggc	gggaccttca	ccggctggag	cggagcctgc	2220
tccggcacct	ccaccacctg	caccgtctcc	atgaccacag	cccggctcgt	caccgctact	2280
tttagcgggc	gtgggtgacac	caggccgagc	gcggggtgtg	gtaagaaccg	gacactgcag	2340
aatggcaca	tcaccatttc	aagtggcggc	gtcaaccgca	cctacatcct	acgcacgcct	2400
gacaactaca	acaacacgca	tgataaccgg	ctgatcatgg	cttatcactg	gcttaacggc	2460
agcgcgcgag	atgtggcgag	cgagaactac	taccggctgt	tcccactctc	caacaacagc	2520
accatcttcg	tgggcgctca	ggggctggat	gccggatggg	ctaacaccaa	caaccgcgac	2580
ctgaacctca	ccgatgccat	actcaccacg	gtcgagaacg	atctgtgcgt	cgacttgaac	2640
cgggtctggg	ccaccgggtt	cagctacggc	gcaggtatgt	catacgccat	cgctgtgcc	2700
agggccaatg	tggttcgggg	cgctcgcttc	tatgcggcg	cgagctcag	cggttgacc	2760
gggtgaacca	cggccattgc	gtacttcgca	acgcacggca	tcaacgacag	tgctctcaac	2820
atctcgcaag	ggcgactct	acgcgaccgc	ttgtctcga	acaacagctg	cacggcgag	2880
aaccctcccg	agccttcctc	gggcagcggg	acgcacatct	gcagctccta	ccagaactgc	2940
tcggcaggac	atcctgtccg	gtggtgcgcg	ttcgacggcg	accacacccc	gaatcagacc	3000
gaccgcggcc	agagcacaag	ctaa				3024

<210> 84
 <211> 1007
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(30)

<400> 84

Met	Lys	Thr	Lys	Arg	Ser	Ile	Phe	Arg	Leu	Ser	Ile	Leu	Val	Val	Leu
1				5					10					15	
Ala	Val	Leu	Leu	Phe	Ser	Ala	Ile	Thr	Leu	Thr	Ala	Ser	Ala	Ala	Asp
			20					25					30		
Thr	Leu	Gly	Ala	Ala	Ala	Ala	Gln	Ser	Gly	Arg	Tyr	Phe	Gly	Thr	Ala
		35					40					45			
Ile	Ala	Ala	Gly	Lys	Leu	Gly	Asp	Ser	Thr	Tyr	Thr	Thr	Ile	Ala	Asn
	50					55					60				
Arg	Glu	Phe	Asn	Met	Ile	Thr	Ala	Glu	Asn	Glu	Met	Lys	Ile	Asp	Ala
65					70				75					80	
Thr	Glu	Pro	Asn	Gln	Asn	Gln	Phe	Asn	Phe	Thr	Asn	Ala	Asp	Arg	Ile
			85						90					95	
Phe	Asn	Trp	Ala	Val	Gln	Asn	Gly	Lys	Gln	Val	Arg	Gly	His	Thr	Leu
			100					105					110		
Ala	Trp	His	Ser	Gln	Gln	Pro	Gly	Trp	Met	Ser	Ser	Met	Ser	Gly	Thr
	115						120					125			
Ala	Leu	Arg	Asn	Ala	Met	Ile	Asn	His	Ile	Asn	Gly	Val	Met	Ala	His
	130					135					140				
Tyr	Lys	Gly	Arg	Ile	Tyr	Ala	Trp	Asp	Val	Val	Asn	Glu	Ala	Phe	Asn
145					150				155					160	
Glu	Asp	Gly	Ser	Arg	Arg	Asn	Ser	Asn	Leu	Gln	Gln	Thr	Gly	Asn	Asp
			165						170					175	
Trp	Ile	Glu	Val	Ala	Phe	Arg	Thr	Ala	Arg	Thr	Ala	Asp	Pro	Ala	Ala
			180					185					190		
Lys	Leu	Cys	Tyr	Asn	Asp	Tyr	Asn	Ile	Glu	Ala	Trp	Ser	Tyr	Ala	Lys
	195						200					205			
Thr	Gln	Gly	Val	Tyr	Arg	Met	Val	Gln	Asp	Phe	Lys	Ser	Arg	Gly	Val
	210					215					220				
Pro	Ile	Asp	Cys	Val	Gly	Phe	Gln	Ser	His	Phe	Asn	Ser	Gly	Thr	Ser
225					230				235					240	
Tyr	Val	Asn	Ser	Asn	Phe	Arg	Thr	Thr	Leu	Gln	Ser	Phe	Ala	Ala	Leu
			245						250					255	
Gly	Val	Asp	Val	Gln	Ile	Thr	Glu	Leu	Asp	Val	Glu	Asn	Ala	Asp	Ser
			260					265					270		
Arg	Leu	Asp	Trp	Trp	Arg	Gly	Ile	Val	Asn	Asp	Cys	Leu	Ala	Val	Pro

Page 69

Leu Phe Pro Leu Ser Asn Asn Ser Thr Ile Phe Val Ala Pro Gln Gly
 835 840 845
 Leu Asp Ala Gly Trp Ala Asn Thr Asn Asn Arg Asp Leu Asn Leu Thr
 850 855 860
 Asp Ala Ile Leu Thr Gln Val Glu Asn Asp Leu Cys Val Asp Leu Asn
 865 870 875 880
 Arg Val Trp Ala Thr Gly Phe Ser Tyr Gly Ala Gly Met Ser Tyr Ala
 885 890 895
 Ile Ala Cys Ala Arg Ala Asn Val Phe Arg Gly Val Ala Leu Tyr Ala
 900 905 910
 Gly Ala Gln Leu Ser Gly Cys Thr Gly Gly Thr Thr Ala Ile Ala Tyr
 915 920 925
 Phe Ala Thr His Gly Ile Asn Asp Ser Val Leu Asn Ile Ser Gln Gly
 930 935 940
 Arg Thr Leu Arg Asp Arg Phe Val Ser Asn Asn Ser Cys Thr Ala Gln
 945 950 955 960
 Asn Pro Pro Glu Pro Ser Ser Gly Ser Gly Thr His Ile Cys Thr Ser
 965 970 975
 Tyr Gln Asn Cys Ser Ala Gly His Pro Val Arg Trp Cys Ala Phe Asp
 980 985 990
 Gly Asp His Thr Pro Asn Gln Thr Asp Arg Gly Gln Ser Thr Ser
 995 1000 1005

<210> 85
 <211> 1254
 <212> DNA
 <213> Bacteria

<400> 85
 atgaccttga ttacgccaa ctcgaaatta accctcacta aagggaacaa aagctggagc 60
 tcgcgcgcct gcaggtcgac actagtggat ctcacacttt acttcgagtc tcagaatccg 120
 accctcgagt tctacgtgga cgatgtgaag gtagtggaca ccacctctgc tgagataaaa 180
 ctcgagatga atccagaaga ggaaatacca gccctcaggg aagttctgaa agactacttc 240
 agagtgggcg ttgctcttcc atccaaggta ttcatacaacc agaaggactt aacgctcatc 300
 accaagcact tcaacagcat caccgcagaa aatgagatga aacctgatag tctgcttgca 360
 ggcattgaga atggcaaaact caagttcaga tttgaaacag cagacaaata catcgaat 420
 gcacagcaaa acggcatggg tgtgaggggc cacacactgg tatggcacia tcagacgccc 480
 gagtggttct tcaaagacga aaatggaaac ctctcttcca aagaagcgat gacagaaaga 540
 ctcagagaat acatacacac cgtcgttgga cacttcaaag ggaaggtcta cgcatgggac 600
 gttgtgaacg aagcgggtcga tccgaaccag ccagatggac tgagaagatc cacctggtat 660
 cagatcatgg ggctgacta catagaactt gccttcaagt ttgcaaggga ggcagatccc 720
 gatgcgaaac tcttctacaa cgactacaac accttcgaac ccaaaaagag agacatcatc 780
 tacaaccttg tgaagagtct caaggaaaag ggtctcatcg atggaatcgg tatgcagtgt 840
 cacatcagtc ttgcaacgga catcaggcag atcgaagagg ccatcaaaaa gttcagctcc 900
 atccctggta tagaaatcca cataacagag ctcgatatga gcgtctacag agattctact 960
 tccaactacc cagaggcacc gaggaacgca ctcatgaaac aggtcacaa gatggctcaa 1020
 ctctttgaaa tcttcaagaa atacagtaat gtgatcacia acgtcacgtt ctggggcttc 1080
 aaagacgact actcctggag agcaacaaga agaaatgact ggacattgat ctttgacaaa 1140
 gattatcagg caaaactcgc ttactgggcg attgctgctc ctgaagtgct accacctctt 1200
 tcaaaagaaa gcaagatcca aagaattcaa aaagcttctc gagagtactt cttag 1254

<210> 86
 <211> 417
 <212> PRT
 <213> Bacteria

<400> 86
 Met Thr Leu Ile Thr Pro Ser Ser Lys Leu Thr Leu Thr Lys Gly Asn
 1 5 10 15
 Lys Ser Trp Ser Ser Arg Ala Cys Arg Ser Thr Leu Val Asp Leu Thr
 20 25 30
 Leu Tyr Phe Glu Ser Gln Asn Pro Thr Leu Glu Phe Tyr Val Asp Asp
 35 40 45
 Val Lys Val Val Asp Thr Thr Ser Ala Glu Ile Lys Leu Glu Met Asn
 50 55 60
 Pro Glu Glu Glu Ile Pro Ala Leu Arg Glu Val Leu Lys Asp Tyr Phe
 65 70 75 80
 Arg Val Gly Val Ala Leu Pro Ser Lys Val Phe Ile Asn Gln Lys Asp
 85 90 95

Leu Thr Leu Ile Thr Lys His Phe Asn Ser Ile Thr Ala Glu Asn Glu
 100 105 110
 Met Lys Pro Asp Ser Leu Leu Ala Gly Ile Glu Asn Gly Lys Leu Lys
 115 120 125
 Phe Arg Phe Glu Thr Ala Asp Lys Tyr Ile Glu Phe Ala Gln Gln Asn
 130 135 140
 Gly Met Val Val Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro
 145 150 155 160
 Glu Trp Phe Phe Lys Asp Glu Asn Gly Asn Leu Leu Ser Lys Glu Ala
 165 170 175
 Met Thr Glu Arg Leu Arg Glu Tyr Ile His Thr Val Val Gly His Phe
 180 185 190
 Lys Gly Lys Val Tyr Ala Trp Asp Val Val Asn Glu Ala Val Asp Pro
 195 200 205
 Asn Gln Pro Asp Gly Leu Arg Arg Ser Thr Trp Tyr Gln Ile Met Gly
 210 215 220
 Pro Asp Tyr Ile Glu Leu Ala Phe Lys Phe Ala Arg Glu Ala Asp Pro
 225 230 235 240
 Asp Ala Lys Leu Phe Tyr Asn Asp Tyr Asn Thr Phe Glu Pro Lys Lys
 245 250 255
 Arg Asp Ile Ile Tyr Asn Leu Val Lys Ser Leu Lys Glu Lys Gly Leu
 260 265 270
 Ile Asp Gly Ile Gly Met Gln Cys His Ile Ser Leu Ala Thr Asp Ile
 275 280 285
 Arg Gln Ile Glu Glu Ala Ile Lys Lys Phe Ser Ser Ile Pro Gly Ile
 290 295 300
 Glu Ile His Ile Thr Glu Leu Asp Met Ser Val Tyr Arg Asp Ser Thr
 305 310 315 320
 Ser Asn Tyr Pro Glu Ala Pro Arg Asn Ala Leu Ile Glu Gln Ala His
 325 330 335
 Lys Met Ala Gln Leu Phe Glu Ile Phe Lys Lys Tyr Ser Asn Val Ile
 340 345 350
 Thr Asn Val Thr Phe Trp Gly Leu Lys Asp Asp Tyr Ser Trp Arg Ala
 355 360 365
 Thr Arg Arg Asn Asp Trp Thr Leu Ile Phe Asp Lys Asp Tyr Gln Ala
 370 375 380
 Lys Leu Ala Tyr Trp Ala Ile Val Ala Pro Glu Val Leu Pro Pro Leu
 385 390 395 400
 Ser Lys Glu Ser Lys Ile Gln Arg Ile Gln Lys Ala Ser Arg Glu Tyr
 405 410 415
 Phe

<210> 87
 <211> 1089
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 87
 ttgaagaaca gaattaaaaa ggttgtgggc gggctcgccc tggcgagtgt tctgctcacc 60
 tcggtaatgg caggcaatgc cagcgagca attaccaatg gatcgaagt cctggggaat 120
 atcattgccg gcagtgtctc aagtaacttc accacctact ggaatcaggt caccgccgag 180
 aacggcacca aatgggggttc catcgaaggc aaccgcaacc agatgaactg gggaaacgcg 240
 gacatgatct ataactacgc catcagcaaa aacatcccgt tcaaattcca tactctcgtc 300
 tggggaagcc agggagccaa ctgggtggcc ggcttgtcgg cagcggagca gaaggcggaa 360
 atcagctcat tcattactca agcaggacag cggtattccg cgaagacagc ttttgtggat 420
 gtagtcaatg aaccgctgca tgccaagcct tcgtaccgca atgccatcgg cggcgatggc 480
 agcaccggct gggattgggt gatctggtct ttccagcaag ccggggccgc cttcccgaac 540
 gccaaagctgc acctcaatga ctacggcatt atcggtgacc ccagcgcggc cgataaatat 600
 gtgaacatta tcaatatcct gaaatccaga ggactgatcg atggtattgg tattcagtgc 660
 cactacttca atatggataa cgtaagtgtg agcaccatga atactgtact gggtaagcct 720
 gctgcaacag gcctgccaat ctatgtctcc gagctggata ttaccggtga tgacaacacc 780
 cagcttgcca gataccaaca gaaattccct gtgctctgga accatccttc cgtgaagggc 840
 gtcaccctgt ggggctacat ccaaaatcag acctgggcat caggcaccga tctggtgaat 900
 tccaacggca cagagcgccc tgccctgaag tggctgaagc aatacctggg cggctcgta 960
 gctctgatgg aaaccacaga cgcccaagac ctactatca ctgacagtct gatccagccg 1020

gacagtgtgg ttgagccgga ccctcaactg gatctccagc cgggtgcttga gcccgttccg 1080
gctgagtaa 1089

<210> 88
<211> 362
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(29)

<400> 88
Leu Lys Asn Arg Ile Lys Lys Val Val Gly Gly Leu Ala Leu Ala Ser
1 5 10 15
Val Leu Leu Thr Ser Val Met Ala Gly Asn Ala Ser Ala Ala Ile Thr
20 25 30
Asn Gly Ser Lys Phe Leu Gly Asn Ile Ile Ala Gly Ser Ala Pro Ser
35 40 45
Asn Phe Thr Thr Tyr Trp Asn Gln Val Thr Pro Glu Asn Gly Thr Lys
50 55 60
Trp Gly Ser Ile Glu Gly Asn Arg Asn Gln Met Asn Trp Gly Asn Ala
65 70 75 80
Asp Met Ile Tyr Asn Tyr Ala Ile Ser Lys Asn Ile Pro Phe Lys Phe
85 90 95
His Thr Leu Val Trp Gly Ser Gln Glu Pro Asn Trp Val Ala Gly Leu
100 105 110
Ser Ala Ala Glu Gln Lys Ala Glu Ile Ser Ser Phe Ile Thr Gln Ala
115 120 125
Gly Gln Arg Tyr Ser Ala Lys Thr Ala Phe Val Asp Val Val Asn Glu
130 135 140
Pro Leu His Ala Lys Pro Ser Tyr Arg Asn Ala Ile Gly Gly Asp Gly
145 150 155 160
Ser Thr Gly Trp Asp Trp Val Ile Trp Ser Phe Gln Gln Ala Arg Ala
165 170 175
Ala Phe Pro Asn Ala Lys Leu His Leu Asn Asp Tyr Gly Ile Ile Gly
180 185 190
Asp Pro Ser Ala Ala Asp Lys Tyr Val Asn Ile Ile Asn Ile Leu Lys
195 200 205
Ser Arg Gly Leu Ile Asp Gly Ile Gly Ile Gln Cys His Tyr Phe Asn
210 215 220
Met Asp Asn Val Ser Val Ser Thr Met Asn Thr Val Leu Gly Lys Leu
225 230 235 240
Ala Ala Thr Gly Leu Pro Ile Tyr Val Ser Glu Leu Asp Ile Thr Gly
245 250 255
Asp Asp Asn Thr Gln Leu Ala Arg Tyr Gln Gln Lys Phe Pro Val Leu
260 265 270
Trp Asn His Pro Ser Val Lys Gly Val Thr Leu Trp Gly Tyr Ile Gln
275 280 285
Asn Gln Thr Trp Ala Ser Gly Thr His Leu Val Asn Ser Asn Gly Thr
290 295 300
Glu Arg Pro Ala Leu Lys Trp Leu Lys Gln Tyr Leu Gly Gly Ser Ser
305 310 315 320
Ala Leu Met Glu Thr Thr Asp Ala Gln Asp Leu Thr Ile Thr Asp Ser
325 330 335
Leu Ile Gln Pro Asp Ser Val Val Glu Pro Asp Pro Gln Leu Asp Leu
340 345 350
Gln Pro Val Leu Glu Pro Val Pro Ala Glu
355 360

<210> 89
<211> 2541
<212> DNA
<213> Bacteria

<400> 89
atggatacat tgttcaatac aaccgatgag cgtggggcgt ccaagcgccg tggcatcgtc 60

```

gcggcgccttg cgccgcagc catgctggtg ccgctggcgt tcgccccgac ggccatggcg 120
gccgaccccg actatccggg cgccatcaag ggcgaataca atccgctggg aatcaacgct 180
ggtgtcgcca tcgagacata caccctcaac caggacaagg agaaggccct ggtcgagaac 240
ttcgaccaga tcaccccgga gaactcgctg aagccggaag gctggtacga cgaccagcat 300
aatctccgca tgtcgatga cgcgcggaac ctgctgacgt tcgccagcga gaacggcatc 360
aaggctctacg gccatgttct ggtctggcac tcgcagacgc ccgactggtt cttccaggcc 420
gacgaatggg gccatgacac caacgacaac ccggcgctca ccagctgccc gcttgccgac 480
aaggccacga tgcaggaacg ccagcgcagg cacatcgaga acgtggcgga ggccatctcc 540
gacgaattcg gaaaattcgg cagcccgacc aatcccgtcg tcgcgttcga cgtggtcaac 600
gagaccgtga acgacagcga cgaccccgcc accaacggca tgcgcaattc gctgtggtat 660
cagacctatg gggcgagga ctacatctat gacgcgttcc ggaacgcgaa tacgtatctg 720
aacgacgtct acgcccggga cgacgcggag catccgggtga cgttggttcat caacgattac 780
ggcaccgagc aggcgggcaa gcggtcccg cagttccatg tgcgttgac cacggcctcg 840
caggggggttc ctttgacgg catcggtcac gacctatg tcctcgctcg gcaagaagca ggccatcacc 900
tcgaatctcg acgacgcgct gaccgatatg atccgctcg accggagcga agctcatcga gcaggagcgg 1020
gaactggacg tcgccaccgg aacgcccgtt acggagcgcc accagctgtt ctcgggttccg 1080
tactactacg acgtcaacca gatcatccac ccagtcctgg cgcaacaagg agggcgcgcc gctgctgttc 1140
gtgtgggggc tgagcgacga gccggcgctac atcggtctaca tcggtgatag cgccaacctt 1200
gacgacaacc tggagaagaa gcccggcggtt aaggtatgac ccgtgggcat cgactcggcg 1260
cccgcggcgt tgaagagcat gaacgcattc gtcggcgcg gtccctccgt ttcgtgtgtc 1320
cttcccggta ccgtggcgga gtcggcgcg ttccggctcg tcaatgtcta ttggaaggac 1380
gagatgaccc cgtctcgta tgacgccgtt gatgccagcg cggcgatga cgacaccgtc 1440
ggctctcttg tcgtctacgc ggatgtcgcc atcgggcgca acggtgtgac cggcgcgag 1500
accgtgcgtg tgggtgacgc cgagtatacg gccggatacg aagtcgtggc cgatatcccg 1560
ggtgtgcagg ccaacgtcgt ttcgtctgat atgaacgtca tcgcgacgga ttccgccacc 1620
tacaccggtg cagagaagga catcgtcgag cacgaacgac actggcgccg tcacgctggc cgagccgctg 1680
acggagacca gcgctggag ggttcccggc gacgcccagg ctccggtcgt tgacgccgac 1740
agctacacgg aagccgtgaa ggaagccaac gaggttcccg tgggtaagg gaccgccgcc 1800
ccgtcggatt ccgtctggg cggaggcgac cgctaccgcc aagaccctgt ggtcggacgg caagctgtat 1860
acgccttccc cggaggcgac cgcgacatc gatctgacca actcgaatcc gtgggagaag 1920
gtcctcatgg aagtgtacat cgaccgtggc aacaccaaga gcggccagta taccaacgac 1980
gactccgttg aggtgtacat cgccgatggt gcggagctga gcttcggctc cggcgcgctg 2040
atccagcaga ttcgctgtc ggtccagacc gccggcaagc tcgtcgatgg cggctatgtc 2100
gaggtatgtc agaagtcctt ccatcgatct gggaaacggc atcggcattc gcaactgggc cgatccgacc 2160
gtcgagatgg cagatcaacg acgcaagaa cgggtgctga ggcgtgctgc gtctgtggc cgatccctcc 2220
ggtgccggct atcagacggc gtcccattgg agaagatccc gagacccccg gtgacgagga gactcctggc 2280
gaaaccgaga ccccgggtgg agacagcctg cgacgaggaa acccccgggt aggataccga gaagcctggc 2340
gaggataacc agaagcctgg cgacgaggaa acccccgggt aggataccga gaagcctggc 2400
gacgagaagc cgcggccttc cgacgatgct gacaacgacg acaagatgcc gcagaccggt 2460
tccgcggtca tcggaatcgc cgtggtggcg ctgctgctgg ttgccgccgg atgcgggctg 2520
gtcatcgctc ggcgtcgatg a

```

<210> 90
 <211> 846
 <212> PRT
 <213> Bacteria

<220>
 <221> SIGNAL
 <222> (1)...(40)

<400> 90
 Met Asp Thr Leu Phe Asn Thr Thr Asp Glu Arg Gly Ala Ser Lys Arg
 1 5 10 15
 Arg Gly Ile Val Ala Ala Leu Ala Ala Ala Met Leu Val Pro Leu
 20 25 30
 Ala Phe Ala Pro Thr Ala Met Ala Ala Asp Pro Asp Tyr Pro Gly Gly
 35 40 45
 Ile Lys Gly Glu Tyr Asn Pro Leu Gly Ile Asn Ala Gly Val Ala Ile
 50 55 60
 Glu Thr Tyr Thr Leu Asn Gln Asp Lys Glu Lys Ala Leu Val Glu Asn
 65 70 75 80
 Phe Asp Gln Ile Thr Pro Glu Asn Ser Leu Lys Pro Glu Gly Trp Tyr
 85 90 95
 Asp Asp Gln His Asn Phe Arg Met Ser Asp Asp Ala Arg Asn Leu Leu
 100 105 110
 Thr Phe Ala Ser Glu Asn Gly Ile Lys Val Tyr Gly His Val Leu Val
 115 120 125

Trp	His	Ser	Gln	Thr	Pro	Asp	Trp	Phe	Phe	Gln	Ala	Asp	Glu	Trp	Cys
130	130					135					140				
His	Asp	Thr	Asn	Asp	Asn	Pro	Gly	Val	Thr	Ser	Cys	Pro	Leu	Ala	Asp
145					150					155					160
Lys	Ala	Thr	Met	Gln	Glu	Arg	Gln	Arg	Arg	His	Ile	Glu	Asn	Val	Ala
			165						170					175	
Glu	Ala	Ile	Ser	Asp	Glu	Phe	Gly	Lys	Phe	Gly	Ser	Pro	Thr	Asn	Pro
			180					185					190		
Val	Val	Ala	Phe	Asp	Val	Val	Asn	Glu	Thr	Val	Asn	Asp	Ser	Asp	Asp
		195					200					205			
Pro	Ala	Thr	Asn	Gly	Met	Arg	Asn	Ser	Leu	Trp	Tyr	Gln	Thr	Tyr	Gly
		210				215					220				
Gly	Glu	Asp	Tyr	Ile	Tyr	Asp	Ala	Phe	Arg	Asn	Ala	Asn	Thr	Tyr	Leu
225					230					235					240
Asn	Asp	Val	Tyr	Ala	Ala	Asp	Asp	Ala	Glu	His	Pro	Val	Thr	Leu	Phe
				245					250					255	
Ile	Asn	Asp	Tyr	Gly	Thr	Glu	Gln	Ala	Gly	Lys	Arg	Ser	Arg	Tyr	Lys
			260					265					270		
Ala	Leu	Leu	Glu	Arg	Met	Ile	Gln	Gln	Gly	Val	Pro	Phe	Asp	Gly	Ile
		275					280					285			
Gly	His	Gln	Phe	His	Val	Ser	Leu	Thr	Thr	Ala	Ser	Ser	Asn	Leu	Asp
	290					295					300				
Asp	Ala	Leu	Thr	Asp	Met	Ser	Ser	Leu	Gly	Lys	Lys	Gln	Ala	Ile	Thr
305					310					315					320
Glu	Leu	Asp	Val	Ala	Thr	Gly	Thr	Pro	Val	Thr	Glu	Ala	Lys	Leu	Ile
				325					330					335	
Glu	Gln	Gly	Arg	Tyr	Tyr	Tyr	Asp	Val	Asn	Gln	Ile	Ile	His	Arg	His
			340					345					350		
Ala	Asp	Gln	Leu	Phe	Ser	Val	Ser	Val	Trp	Gly	Leu	Ser	Asp	Asp	Gln
		355					360					365			
Ser	Trp	Arg	Asn	Lys	Glu	Gly	Ala	Pro	Leu	Leu	Phe	Asp	Asp	Asn	Leu
	370					375					380				
Glu	Lys	Lys	Pro	Ala	Tyr	Ile	Gly	Tyr	Ile	Gly	Asp	Ser	Ala	Asn	Leu
385					390					395					400
Pro	Glu	Pro	Leu	Lys	Ser	Met	Asn	Ala	Phe	Lys	Asp	Asp	Ala	Val	Gly
				405					410					415	
Ile	Asp	Ser	Ala	Leu	Pro	Gly	Thr	Val	Ala	Glu	Ser	Gly	Ala	Ser	Ser
			420					425					430		
Pro	Trp	Glu	Arg	Leu	Ser	Leu	Val	Glu	Met	Thr	Pro	Ser	Ala	Tyr	Asp
		435					440					445			
Ala	Val	Ser	Gly	Ser	Phe	Asn	Val	Tyr	Trp	Lys	Asp	Gly	Ser	Leu	Val
	450					455					460				
Val	Tyr	Ala	Asp	Val	Ala	Asp	Ala	Ser	Ala	Ala	Asp	Asp	Asp	Thr	Val
465					470					475					480
Thr	Val	Arg	Val	Gly	Asp	Ala	Glu	Tyr	Thr	Ile	Gly	Arg	Asn	Gly	Val
				485					490					495	
Thr	Gly	Gly	Glu	Gly	Val	Gln	Ala	Asn	Val	Val	Ser	Ser	Asp	Ala	Gly
			500					505					510		
Tyr	Glu	Val	Val	Ala	Asp	Ile	Pro	Tyr	Thr	Gly	Ala	Glu	Lys	Asp	Ile
	515						520					525			
Val	Glu	Met	Asn	Val	Ile	Ala	Thr	Asp	Ser	Ala	Thr	Thr	Glu	Thr	Ser
	530					535					540				
Ala	Trp	Ser	Thr	Asn	Asp	Thr	Gly	Ala	Val	Thr	Leu	Ala	Glu	Pro	Leu
545					550					555					560
Ser	Tyr	Thr	Glu	Ala	Val	Lys	Val	Pro	Ala	Asp	Ala	Gln	Ala	Pro	Val
				565					570					575	
Val	Asp	Ala	Asp	Pro	Ser	Asp	Ser	Val	Trp	Ala	Glu	Ala	Asn	Glu	Val
			580					585					590		
Pro	Val	Gly	Lys	Val	Thr	Ala	Ala	Thr	Pro	Ser	Pro	Glu	Ala	Thr	Ala
		595					600					605			
Thr	Ala	Lys	Thr	Leu	Trp	Ser	Asp	Gly	Lys	Leu	Tyr	Val	Leu	Met	Glu
	610					615					620				
Val	Thr	Asp	Ala	Asp	Ile	Asp	Leu	Thr	Asn	Ser	Asn	Pro	Trp	Glu	Lys
625					630					635					640
Asp	Ser	Val	Glu	Val	Tyr	Ile	Asp	Arg	Gly	Asn	Thr	Lys	Ser	Gly	Gln
				645					650					655	
Tyr	Thr	Asn	Asp	Ile	Gln	Gln	Ile	Arg	Val	Ser	Ala	Asp	Gly	Ala	Glu
			660					665					670		
Leu	Ser	Phe	Gly	Ser	Gly	Ala	Ser	Glu	Asp	Val	Gln	Lys	Ser	Met	Val

675 680 685
 Gln Thr Ala Gly Lys Leu Val Asp Gly Gly Tyr Val Val Glu Met Ala
 690 695 700
 Ile Asp Leu Gly Thr Ala Glu Ala Gly Thr Phe Glu Gly Val Asp Phe
 705 710 715 720
 Gln Ile Asn Asp Ala Lys Asn Gly Ala Arg Ile Gly Ile Arg Asn Trp
 725 730 735
 Ala Asp Pro Thr Gly Ala Gly Tyr Gln Thr Ala Ser His Trp Gly Val
 740 745 750
 Leu Arg Leu Leu Ala Asp Pro Ser Glu Thr Glu Thr Pro Gly Gly Glu
 755 760 765
 Asp Pro Glu Thr Pro Gly Asp Glu Glu Thr Pro Gly Glu Asp Thr Glu
 770 775 780
 Lys Pro Gly Asp Glu Glu Thr Pro Gly Glu Asp Thr Glu Lys Pro Gly
 785 790 795 800
 Asp Glu Lys Pro Arg Pro Ser Asp Asp Ala Asp Asn Asp Asp Lys Met
 805 810 815
 Pro Gln Thr Gly Ser Ala Val Ile Gly Ile Ala Val Val Ala Leu Leu
 820 825 830
 Leu Val Ala Ala Gly Cys Gly Leu Val Ile Ala Arg Arg Arg
 835 840 845

<210> 91
 <211> 1023
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 91
 atgaatgtat cggtagcggc cgagtcgcga cttaaagaca tcttcgcgga agacttccat 60
 ataggtgagg cggtagcagg taatacagatc aagtcgcagg agagtctgct tacgcatcac 120
 tttacacagc ttacggcggg aaacgaaatg aagttcgcca gcgtccatcc agaggaagag 180
 ctttacacct tcgaggaagc ggatcagatc gtggacttcg cgcgcaaaca cgggatggct 240
 gtccgcggac atacgctggt atggcataac cagaccaccg attggttgtt ccgcgacaag 300
 cagaatcagc tcgtgagcaa agccgtgctt tatgaaagaa tccgttcgca tatccaaacg 360
 gtagtaggca gatataaggg cgatatattac gcttgggacg ttgtgaacga ggatcattgcc 420
 gatgacggcg atcagttgct gcgtacctcc agctggacgg aaatcgccgg ggacgaattc 480
 atcgccaaag cgtttgaata cgcgcatgct gccgaccgga atgcgctggt gttctacaac 540
 gactacaatg agtcccatcc aagcaaacgg gataaaattt ataccttggg caagtctctt 600
 ctggaccggg gaggacctat tcacggcatt ggcctgcagg cacactggaa tctgttcaac 660
 ccgtccttgg atgacatccg ggcagccatc gaaaaatatg cttcgtagg attgcagctc 720
 cagctcacgg aactggatgt gtcggtattc cgtttcgaag ataagcggg cgatctgacc 780
 gagcctgaac cgggaatgct ggaacagcag gctgaattct acgaagcgt gttcaagctg 840
 cttaaggaat acagcgatgt aattagcgcg gtgacgttct ggggagctgc ggacgaccac 900
 acctggctca gcgattttcc ggtacgtggg cgcaaaaact ggccgctgct gttcgtatgag 960
 cggcacaggc cgaagccggc atattatcgc ttagctgctc ttgccaatca tcttcggcgt 1020
 tga 1023

<210> 92
 <211> 340
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 92
 Met Asn Val Ser Val Pro Ala Glu Ser Ala Leu Lys Asp Ile Phe Ala
 1 5 10 15
 Glu Asp Phe His Ile Gly Ala Ala Val Ser Ser Asn Thr Ile Lys Ser
 20 25 30
 Gln Glu Ser Leu Leu Thr His His Phe Asn Ser Ile Thr Ala Glu Asn
 35 40 45
 Glu Met Lys Phe Ala Ser Val His Pro Glu Glu Glu Leu Tyr Thr Phe
 50 55 60
 Glu Glu Ala Asp Gln Ile Val Asp Phe Ala Arg Lys His Gly Met Ala
 65 70 75 80

Val Arg Gly His Thr Leu Val Trp His Asn Gln Thr Thr Asp Trp Leu
 85 90 95
 Phe Arg Asp Lys Gln Asn Gln Leu Val Ser Lys Ala Val Leu Tyr Glu
 100 105 110
 Arg Ile Arg Ser His Ile Gln Thr Val Val Gly Arg Tyr Lys Gly Asp
 115 120 125
 Ile Tyr Ala Trp Asp Val Val Asn Glu Val Ile Ala Asp Asp Gly Asp
 130 135 140
 Gln Leu Leu Arg Thr Ser Ser Trp Thr Glu Ile Ala Gly Asp Glu Phe
 145 150 155 160
 Ile Ala Lys Ala Phe Glu Tyr Ala His Ala Ala Asp Pro Asn Ala Leu
 165 170 175
 Leu Phe Tyr Asn Asp Tyr Asn Glu Ser His Pro Ser Lys Arg Asp Lys
 180 185 190
 Ile Tyr Thr Leu Val Lys Ser Leu Leu Asp Arg Gly Val Pro Ile His
 195 200 205
 Gly Ile Gly Leu Gln Ala His Trp Asn Leu Phe Asn Pro Ser Leu Asp
 210 215 220
 Asp Ile Arg Ala Ala Ile Glu Lys Tyr Ala Ser Leu Gly Leu Gln Leu
 225 230 235 240
 Gln Leu Thr Glu Leu Asp Val Ser Val Phe Arg Phe Glu Asp Lys Arg
 245 250 255
 Ala Asp Leu Thr Glu Pro Glu Pro Gly Met Leu Glu Gln Gln Ala Glu
 260 265 270
 Phe Tyr Glu Ala Val Phe Lys Leu Leu Lys Glu Tyr Ser Asp Val Ile
 275 280 285
 Ser Ala Val Thr Phe Trp Gly Ala Ala Asp Asp His Thr Trp Leu Ser
 290 295 300
 Asp Phe Pro Val Arg Gly Arg Lys Asn Trp Pro Leu Leu Phe Asp Glu
 305 310 315 320
 Arg His Arg Pro Lys Pro Ala Tyr Tyr Arg Leu Ala Ala Leu Ala Asn
 325 330 335
 His Leu Arg Arg
 340

<210> 93
 <211> 1011
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 93
 atgaatcaat cagtaaata agcacagggt cctgcattat cggatgtata tgaagattat 60
 ttttcaatag gtgccgctgt taatccactt acttttaggta cgcaaaaaaa gctgttaacc 120
 aaacatttta atagtataac ggctgagaat gaaatgaaat ttgaagcatt acagcctaaa 180
 ccagatcaat ttacatttga tacggcggat aaaatgggtg cctttgcccc agcacatgat 240
 atgaagatgc gtggccatac attaatctgg cacaatcaaa caccagattg gatgtttttg 300
 caaaaagacg gtacgacaat tgatcgtgaa acactcttgg agagaatgaa aaaacatatt 360
 aagacgggtg tggaaagata taaaggcaaa atatattgtt gggacgttgt aaatgaagcg 420
 gtagctgatg aaggcgaagc tattttaaga ccatcaaaat ggacggacat tattggcgac 480
 tcgtttattg agtatgcttt taaatacgcc caccgagccg atcccgatgc actgttgttt 540
 tacaatgact acaatgcttg ccaccctcat aaaagagata agatttatca acttgtaaag 600
 gggttaatag acaagggtgt gcccatacac ggtattggcc tacaagcaca ttggaacatt 660
 gttgaccctg cttacgatga tattaaacga gccatcgaaa cttatgcatc attaggatta 720
 agcatacact ttactgaaat ggatgtgtct gtttttgaat atcatgatcg aagaacagac 780
 ttattggaac ctacaaaaga tatggtttca cgtcaagctg agcgttatca ggcatttttt 840
 gaaatatttta ggtcgtatgc tgatgtgatt gattccgtta cgttttgggg catggccgat 900
 gattatacat ggcttgatga ttttccggtg acagggtcgaa aaaattggcc ctttgtattt 960
 gatgcgagac atcagcctaa aacagcattc tggaacatcg ttgattttta a 1011

<210> 94
 <211> 336
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample
 Page 76

<400> 94
 Met Asn Gln Ser Val Asn Glu Ala Gln Val Pro Ala Leu Ser Asp Val
 1 5 10 15
 Tyr Glu Asp Tyr Phe Ser Ile Gly Ala Ala Val Asn Pro Leu Thr Leu
 20 25 30
 Gly Thr Gln Lys Lys Leu Leu Thr Lys His Phe Asn Ser Ile Thr Ala
 35 40 45
 Glu Asn Glu Met Lys Phe Glu Ala Leu Gln Pro Lys Pro Asp Gln Phe
 50 55 60
 Thr Phe Asp Thr Ala Asp Lys Met Val Ala Phe Ala Gln Ala His Asp
 65 70 75 80
 Met Lys Met Arg Gly His Thr Leu Ile Trp His Asn Gln Thr Pro Asp
 85 90 95
 Trp Met Phe Leu Gln Lys Asp Gly Thr Thr Ile Asp Arg Glu Thr Leu
 100 105 110
 Leu Glu Arg Met Lys Lys His Ile Lys Thr Val Val Glu Arg Tyr Lys
 115 120 125
 Gly Lys Ile Tyr Cys Trp Asp Val Val Asn Glu Ala Val Ala Asp Glu
 130 135 140
 Gly Glu Ala Ile Leu Arg Pro Ser Lys Trp Thr Asp Ile Ile Gly Asp
 145 150 155 160
 Ser Phe Ile Glu Tyr Ala Phe Lys Tyr Ala His Glu Ala Asp Pro Asp
 165 170 175
 Ala Leu Leu Phe Tyr Asn Asp Tyr Asn Ala Cys His Pro His Lys Arg
 180 185 190
 Asp Lys Ile Tyr Gln Leu Val Lys Gly Leu Ile Asp Lys Gly Val Pro
 195 200 205
 Ile His Gly Ile Gly Leu Gln Ala His Trp Asn Ile Val Asp Pro Ser
 210 215 220
 Tyr Asp Asp Ile Lys Arg Ala Ile Glu Thr Tyr Ala Ser Leu Gly Leu
 225 230 235 240
 Ser Ile His Phe Thr Glu Met Asp Val Ser Val Phe Glu Tyr His Asp
 245 250 255
 Arg Arg Thr Asp Leu Leu Glu Pro Thr Lys Asp Met Val Ser Arg Gln
 260 265 270
 Ala Glu Arg Tyr Gln Ala Phe Phe Glu Ile Phe Arg Ser Tyr Ala Asp
 275 280 285
 Val Ile Asp Ser Val Thr Phe Trp Gly Met Ala Asp Asp Tyr Thr Trp
 290 295 300
 Leu Asp Asp Phe Pro Val Thr Gly Arg Lys Asn Trp Pro Phe Val Phe
 305 310 315 320
 Asp Ala Arg His Gln Pro Lys Thr Ala Phe Trp Asn Ile Val Asp Phe
 325 330 335

<210> 95
 <211> 1143
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 95
 atgaaaaaaa cgattgcaca ttccacctta tggatagtgt tttttctctt cacttcctgt 60
 gctgttacgg cgcagaagaa tgctaagaat acaagagtaa aactcactac cctaaaagag 120
 gcttaccagg gtaaatctta tatcggtact gcgatgaatc tgagacagat tcacggagat 180
 gatccccagt ctgaaaatat tatcaaaaaa cagttcaatt ccatagtgtc cgaaaactgc 240
 atgaagagta tgtatcttca gccggaggaa ggaaaatttt tcttcgatga tgcggacaag 300
 tttgtggatt ttggtcttca gaacaatatg ttcattcatcg ggcatgtgtt gatttggcat 360
 tcgcaggcgc caaaatgggt ttccaccgat gagaatggaa aaacgggttc cccagaagtt 420
 cttaaacaaa ggatgaaagc ccatatcacc gctgtcgttt cccgctacaa agggaaaatc 480
 aaaggttggg atgtggtgaa cgaagccatt atggaagatg gttcttaccg caaaagcaaa 540
 ttttatgaga ttttgggaga agaatttatt ccgttggcat ttcagtatgc gcatgaagca 600
 gatcctgatg cagaacttta ttacaacgat tataacgaat ggatccccgg aaaaagagct 660
 acggtgacca agataatccg cgatttcaaa tctagaggaa tccgcattga tgccatcggg 720
 atgcaggctc atttcgggat ggattcgccc actttagaag agtatgaaca aaccattcag 780
 ggctatataa aagaaggcgt gaaagtcaat attacggaac tcgatttgag tccgcttcct 840
 tctccttggg gaacttccgc caatgttgcc gatacgagc agtatcagga aaaaatgaat 900

ccttacacca	aaggacttcc	cgccgatgtg	gaaaaagcat	gggaaaaccg	ctatctcgat	960
tttttcaaac	tggttcctgaa	atatcatcag	catatcgagc	gtgttacgtt	ttggggcggt	1020
agcgatatcg	attcctggaa	gaacgatttt	ccagtaagag	gacgtaccga	ttatccacta	1080
ccgtttaacc	gacagtatca	ggcaaaacct	ttggtgcaga	aattaataga	cttaacgaaa	1140
tag						1143

<210> 96
 <211> 380
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(24)

<400> 96

Met	Lys	Lys	Thr	Ile	Ala	His	Phe	Thr	Leu	Trp	Ile	Val	Phe	Phe	Leu
1				5					10				15		
Phe	Thr	Ser	Cys	Ala	Val	Thr	Ala	Gln	Lys	Asn	Ala	Lys	Asn	Thr	Arg
			20					25					30		
Val	Lys	Leu	Thr	Thr	Leu	Lys	Glu	Ala	Tyr	Gln	Gly	Lys	Phe	Tyr	Ile
		35					40					45			
Gly	Thr	Ala	Met	Asn	Leu	Arg	Gln	Ile	His	Gly	Asp	Asp	Pro	Gln	Ser
	50					55					60				
Glu	Asn	Ile	Ile	Lys	Lys	Gln	Phe	Asn	Ser	Ile	Val	Ala	Glu	Asn	Cys
65					70					75					80
Met	Lys	Ser	Met	Tyr	Leu	Gln	Pro	Glu	Glu	Gly	Lys	Phe	Phe	Phe	Asp
				85					90					95	
Asp	Ala	Asp	Lys	Phe	Val	Asp	Phe	Gly	Leu	Gln	Asn	Asn	Met	Phe	Ile
			100					105					110		
Ile	Gly	His	Cys	Leu	Ile	Trp	His	Ser	Gln	Ala	Pro	Lys	Trp	Phe	Phe
		115					120					125			
Thr	Asp	Glu	Asn	Gly	Lys	Thr	Val	Ser	Pro	Glu	Val	Leu	Lys	Gln	Arg
	130					135					140				
Met	Lys	Ala	His	Ile	Thr	Ala	Val	Val	Ser	Arg	Tyr	Lys	Gly	Lys	Ile
145					150					155					160
Lys	Gly	Trp	Asp	Val	Val	Asn	Glu	Ala	Ile	Met	Glu	Asp	Gly	Ser	Tyr
				165					170					175	
Arg	Lys	Ser	Lys	Phe	Tyr	Glu	Ile	Leu	Gly	Glu	Glu	Phe	Ile	Pro	Leu
			180					185					190		
Ala	Phe	Gln	Tyr	Ala	His	Glu	Ala	Asp	Pro	Asp	Ala	Glu	Leu	Tyr	Tyr
		195					200					205			
Asn	Asp	Tyr	Asn	Glu	Trp	Tyr	Pro	Gly	Lys	Arg	Ala	Thr	Val	Thr	Lys
		210				215					220				
Ile	Ile	Arg	Asp	Phe	Lys	Ser	Arg	Gly	Ile	Arg	Ile	Asp	Ala	Ile	Gly
225					230					235					240
Met	Gln	Ala	His	Phe	Gly	Met	Asp	Ser	Pro	Thr	Leu	Glu	Glu	Tyr	Glu
			245						250					255	
Gln	Thr	Ile	Gln	Gly	Tyr	Ile	Lys	Glu	Gly	Val	Lys	Val	Asn	Ile	Thr
			260					265					270		
Glu	Leu	Asp	Leu	Ser	Pro	Leu	Pro	Ser	Pro	Trp	Gly	Thr	Ser	Ala	Asn
		275					280					285			
Val	Ala	Asp	Thr	Gln	Gln	Tyr	Gln	Glu	Lys	Met	Asn	Pro	Tyr	Thr	Lys
	290					295					300				
Gly	Leu	Pro	Ala	Asp	Val	Glu	Lys	Ala	Trp	Glu	Asn	Arg	Tyr	Leu	Asp
305					310					315					320
Phe	Phe	Lys	Leu	Phe	Leu	Lys	Tyr	His	Gln	His	Ile	Glu	Arg	Val	Thr
				325					330					335	
Phe	Trp	Gly	Val	Ser	Asp	Ile	Asp	Ser	Trp	Lys	Asn	Asp	Phe	Pro	Val
			340					345					350		
Arg	Gly	Arg	Thr	Asp	Tyr	Pro	Leu	Pro	Phe	Asn	Arg	Gln	Tyr	Gln	Ala
		355					360					365			
Lys	Pro	Leu	Val	Gln	Lys	Leu	Ile	Asp	Leu	Thr	Lys				
	370					375					380				

<210> 97
 <211> 1407

<212> DNA
<213> Unknown

<220>
<223> obtained from an environmental sample

<400> 97
 atgaatgaaa cctcgcggaa ttggttggag agaggattgc ctttcgaacg ccaacggcgt 60
 tccaacattc agcccagggt tggcgcttgc gcctaccctg ggttggaagc aatcgctcca 120
 tcaaccctga aagggttgca gcggagggtt gcacaagacc gatacaacc tttcaggatt 180
 ggctttctcc cttttccacc cagggtagcg cctgcggcgc aaccctgggc tgatggatca 240
 gaacgcccgtt ggcgttccc gaaacctgcg aagaaacaac tcgccttcct ggccatcacc 300
 agtctcctct cgggtctgct gtggggcgcc gaagtgaac cggcactgaa agacgtattc 360
 cgccaggact tcctgctggg ggcggcggtt aacgcggagc aggtgctgga caccaaccgg 420
 gtcgagtcgg tattgatcga aaagcatttc aacacgatca cgcccagaaa tgtgctgaag 480
 tgggaacgag tccatcctca gcccaccag tattcttttg aggacgcgga tcgctacgtc 540
 gagttcggcc gcaaacacgg aatggtcac atcgccaca cgctggctcg gcacagccag 600
 acgcccggct gggctctccg ggaatgccgac ggaaagacgc tgacgcgcga agccctgctg 660
 gagcggatgc gcgaccacat ccacaccgtg gtcgggcgct acaagggcaa gatccgcggc 720
 tgggatgttg tgaacgaggc gctgcgcgac gacggcgctt ggcgggaatt ccaatggcgg 780
 cggatcatcg gcgacgatta cattttgaaa gccttccagt atgccatga ggccgatccg 840
 gatgcggagc tctattacaa cgattattcg ctggagaagc cggccaagcg caatggcgcc 900
 gtggacctgg tgaagcagct ccaggccggc ggggcgaagc tggccggcgt cggcttgacg 960
 gggcactaca acctcgactg gccggagacc gccgagatcg aaaacaccat cgcggcgttc 1020
 gcggagctgg ggctcaagg gatgatcac gagctggacg tcaacgcgct gccgacgccc 1080
 ggccagtcgg gcgaagccga tgtagggatg acgttcggcg gcaatttcgg cggcgataaa 1140
 tggaatcctt tcacgaacgg actgccggcc gcagtggagc aacgcctcgc ggaccgctac 1200
 gctgaaatct tcaggatctt cacgaagcac agccgtcgga tttcgcgct cactttctgg 1260
 ggcgtcaccg accggacctc ctggctcaac aattttccca tccgcggccg gaccaattac 1320
 ccgttgctct ttgatcggg tggggagccc aaaccgcgt tccgatccgt cgtggcggtc 1380
 cgtcagccgc gccagcccgt cgaatga 1407

<210> 98
<211> 468
<212> PRT
<213> Unknown

<220>
<223> obtained from an environmental sample

<400> 98
 Met Asn Glu Thr Ser Arg Asn Trp Leu Glu Arg Gly Leu Pro Phe Glu
 1 5 10 15
 Arg Gln Arg Arg Ser Asn Ile Gln Pro Arg Val Gly Ala Cys Ala Tyr
 20 25 30
 Pro Gly Leu Glu Ala Ile Ala Pro Ser Thr Leu Lys Gly Leu Gln Arg
 35 40 45
 Arg Phe Ala Gln Asp Arg Tyr Asn Pro Phe Arg Ile Gly Phe Leu Pro
 50 55 60
 Phe Pro Pro Arg Val Ala Pro Ala Ala Gln Pro Trp Ala Asp Gly Ser
 65 70 75 80
 Glu Arg Arg Trp Arg Ser Arg Lys Pro Ala Lys Lys Gln Leu Ala Phe
 85 90 95
 Leu Ala Ile Thr Ser Leu Leu Ser Gly Leu Leu Trp Gly Ala Glu Val
 100 105 110
 Gln Pro Ala Leu Lys Asp Val Phe Arg Gln Asp Phe Leu Leu Gly Ala
 115 120 125
 Ala Leu Asn Ala Glu Gln Val Leu Asp Thr Asn Arg Val Glu Ser Val
 130 135 140
 Leu Ile Glu Lys His Phe Asn Thr Ile Thr Pro Glu Asn Val Leu Lys
 145 150 155 160
 Trp Glu Arg Val His Pro Gln Pro Asn Gln Tyr Ser Phe Glu Asp Ala
 165 170 175
 Asp Arg Tyr Val Glu Phe Gly Arg Lys His Gly Met Val Ile Ile Gly
 180 185 190
 His Thr Leu Val Trp His Ser Gln Thr Pro Gly Trp Val Phe Arg Asp
 195 200 205
 Ala Asp Gly Lys Thr Leu Thr Arg Glu Ala Leu Leu Glu Arg Met Arg
 210 215 220

Asp His Ile His Thr Val Val Gly Arg Tyr Lys Gly Lys Ile Arg Gly
 225 230 235 240
 Trp Asp Val Val Asn Glu Ala Leu Arg Asp Asp Gly Ala Trp Arg Asn
 245 250 255
 Ser Gln Trp Arg Arg Ile Ile Gly Asp Asp Tyr Ile Leu Lys Ala Phe
 260 265 270
 Gln Tyr Ala His Glu Ala Asp Pro Ala Glu Leu Tyr Tyr Asn Asp
 275 280 285
 Tyr Ser Leu Glu Lys Pro Ala Lys Arg Asn Gly Ala Val Asp Leu Val
 290 295 300
 Lys Gln Leu Gln Ala Gly Gly Ala Lys Leu Ala Gly Val Gly Leu Gln
 305 310 315 320
 Gly His Tyr Asn Leu Asp Trp Pro Glu Thr Ala Glu Ile Glu Asn Thr
 325 330 335
 Ile Ala Ala Phe Ala Glu Leu Gly Leu Lys Val Met Ile Thr Glu Leu
 340 345 350
 Asp Val Asn Ala Leu Pro Thr Pro Gly Gln Ser Gly Glu Ala Asp Val
 355 360 365
 Gly Met Thr Phe Gly Gly Asn Phe Gly Gly Asp Lys Trp Asn Pro Phe
 370 375 380
 Thr Asn Gly Leu Pro Ala Ala Val Glu Gln Arg Leu Ala Asp Arg Tyr
 385 390 395 400
 Ala Glu Ile Phe Arg Ile Phe Thr Lys His Ser Arg Arg Ile Ser Arg
 405 410 415
 Val Thr Phe Trp Gly Val Thr Asp Arg Thr Ser Trp Leu Asn Asn Phe
 420 425 430
 Pro Ile Arg Gly Arg Thr Asn Tyr Pro Leu Leu Phe Asp Arg Ala Gly
 435 440 445
 Glu Pro Lys Pro Ala Phe Arg Ser Val Val Ala Val Arg Gln Pro Arg
 450 455 460
 Gln Pro Val Glu
 465

<210> 99
 <211> 1074
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 99
 gtgcgtcaa gagctagcgc gtactggttc ggcgtggggt tgggtggtggc gctgagcctg 60
 gctcagaccc cttcccccca gtccctgcgc gcgctggccg agcgccaggg gctgctggtg 120
 ggagccgcgg tggacctagc ggccctgtac gacccctcgc agcccagagta cgcccaactc 180
 ctgcgccgcg agttcaacct ggtggtggcc gagaacgcc tgaagtgggc ctccctgagc 240
 aacgcgcggg ggcagtacag cttcacggc gctgacgccc tgggtgcgctt cgcccgccag 300
 cacggccagc gcttgccgcg ccacaccctc atctggcac agcaactgcc cgcgtgggtg 360
 cgcagcggca cttctctccg cgaggccatg ctggcggtga tgcaggagca cattcaggcg 420
 gtggccgggc acttccgcgg ccagggtggc tactgggacg tgggtcaacga ggcggtgagt 480
 gaccggggcg gcctgcgcga gaccccttt ctgcggggcg tgggccccga ctacctcgag 540
 cacgccttcc gcttcgcccg cgccgcccac cccagggcca agctcttcta caacgactac 600
 ggcgcccagc gcatggggcg taaatcggac gagatctacg ccttgctcaa agcgctcaag 660
 gccaaagggg taccgctcga cggggtgggc ttccaggccc acctcgacag caccttctcg 720
 gtccagcagg cgcggatgcg ggagaaccta gagacgcttc gccgacctgg gcctcgaggt 780
 gcacatcacc gagctggacg tgcagtaaa aggggcgggc tcgcgggagg aacggctgga 840
 ggccgaggcc cggatctacg ccgaggtgct ggcgacctgc cgcgcggtcc gcggctgcag 900
 cgccgtgacg ctgtggggct tcaccgacgc ccactcctgg cgagccgcgc ccgaaccctc 960
 gatcttcgac gcgctctacc ggcccaaacc ggcgtaccag gctctgctgc gggctctggg 1020
 aggcaaccct tgagcctttt cagcccagtt ttgccaacga ggacagcact atga 1074

<210> 100
 <211> 357
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(33)

<400> 100
 Val Arg Ser Arg Ala Ser Ala Tyr Trp Phe Gly Val Gly Leu Val Val
 1 5 10 15
 Ala Leu Ser Leu Ala Gln Thr Pro Ser Pro Gln Ser Leu Arg Ala Leu
 20 25 30
 Ala Glu Arg Gln Gly Leu Leu Val Gly Ala Ala Val Asp Leu Ala Ala
 35 40 45
 Leu Tyr Asp Pro Leu Glu Pro Glu Tyr Ala Gln Leu Ala Arg Glu
 50 55 60
 Phe Asn Leu Val Val Ala Glu Asn Ala Met Lys Trp Ala Ser Leu Ser
 65 70 75 80
 Asn Ala Arg Gly Gln Tyr Ser Phe Thr Gly Ala Asp Ala Leu Val Arg
 85 90 95
 Phe Ala Arg Gln His Gly Gln Arg Leu Arg Gly His Thr Leu Ile Trp
 100 105 110
 His Glu Gln Leu Pro Ala Trp Val Arg Ser Gly Thr Phe Ser Arg Glu
 115 120 125
 Ala Met Leu Ala Val Met Gln Glu His Ile Gln Ala Val Ala Gly His
 130 135 140
 Phe Arg Gly Gln Val Ala Tyr Trp Asp Val Val Asn Glu Ala Val Ser
 145 150 155 160
 Asp Arg Gly Gly Leu Arg Glu Thr Pro Phe Leu Arg Ala Val Gly Pro
 165 170 175
 Asp Tyr Leu Glu His Ala Phe Arg Phe Ala Arg Ala Ala Asp Pro Gln
 180 185 190
 Ala Lys Leu Phe Tyr Asn Asp Tyr Gly Ala Asp Gly Met Gly Ala Lys
 195 200 205
 Ser Asp Glu Ile Tyr Ala Leu Leu Lys Ala Leu Lys Ala Lys Gly Val
 210 215 220
 Pro Val Asp Gly Val Gly Phe Gln Ala His Leu Asp Ser Thr Phe Ser
 225 230 235 240
 Val Gln Gln Ala Arg Met Arg Glu Asn Leu Glu Thr Leu Arg Arg Pro
 245 250 255
 Gly Pro Arg Gly Ala His His Arg Ala Gly Arg Ala Ala Lys Arg Gly
 260 265 270
 Gly Leu Ala Gly Gly Thr Ala Gly Ala Gly Pro Asp Leu Arg Arg
 275 280 285
 Gly Ala Gly Asp Leu Pro Arg Gly Pro Arg Leu Gln Arg Arg Asp Ala
 290 295 300
 Val Gly Leu His Arg Arg Pro Leu Leu Ala Ser Arg Arg Arg Thr Pro
 305 310 315 320
 Asp Leu Arg Arg Ala Leu Pro Ala Gln Thr Gly Val Pro Gly Ser Ala
 325 330 335
 Ala Gly Ser Gly Arg Gln Pro Leu Ser Leu Phe Ser Pro Val Leu Pro
 340 345 350
 Thr Arg Thr Ala Leu
 355

<210> 101
 <211> 1131
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 101
 atgaagtatt ggcttacaac cctgggtttta atgatagcgg gaataccctt ggcttttggt 60
 tcttcagcaa agcaagataa atcaaagagt ttgaaagatg ctttcaaaaa caaattctat 120
 atcgggtgtgg ctttgaaccg gagtcaatat ctggaacaaa acgaacaggc ggataaagag 180
 ataaaggcac agttcagctc tattgtagct gagaactgca tgaaaagcga aaatctggaa 240
 cctaaagagg gaaaattctt ctttgacgat gccgatcgtt ttgtcgcttt tggagaaaaa 300
 aatggaatgt acatcattgg acatacctta atttggcatt ctcaagtgcc aaaatggttt 360
 ttcatagata atgaaggcaa agttgtttcc cgggaagtgt tgattgaacg aatgaaaaac 420
 tacatccata cagttgtcgg tcattataaa ggtcagagta aaggttggga tgttgtcaat 480
 gaggccattc tagatgatgg ctcatttaga caaagtaatt tctttaaaat actaggagcc 540

gattttatta	aacttgcttt	tcaatttgcc	catgaagcag	atcccaatgc	tgagctttat	600
tacaacgatt	attcgatgtc	caatccgacc	aaaagagacg	gagtgggttcg	catggtgaag	660
tcattgcagc	aacaaggtgt	gagaatagac	gctatcggaa	tgacgggaca	cgtagggatg	720
gattatccca	agttggatga	gtttgaaaat	agtatcaaag	ctttttcgtc	tttaggaacc	780
aaagtgatga	ttacggaact	cgatttaagt	gtcctaccaa	ctcctaaagg	aaaacaaggt	840
gctaataatt	cggatgttgc	cgcttatgag	gaaaagataa	atccttacaa	aaatggctcg	900
ccggctgaag	ttgaaaaggc	ttgggaagac	cggtatttgg	attttttcaa	attatttttg	960
aaatatcaac	accaaatttc	aagggttaca	ttatgggggc	ttagtgatca	ggattcgtgg	1020
aaaaatgatt	tcccagtcag	agggagaacg	gattatcctt	tgcttttcga	cagacaatac	1080
aaaccaaacc	ctgtagtica	gaaaattatt	aaattagcat	tgaaaaaata	a	1131

<210> 102

<211> 376

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(23)

<400> 102

Met	Lys	Tyr	Trp	Leu	Thr	Thr	Leu	Val	Leu	Met	Ile	Ala	Gly	Ile	Pro
1				5				10						15	
Leu	Ala	Phe	Gly	Ser	Ser	Ala	Lys	Gln	Asp	Lys	Ser	Lys	Ser	Leu	Lys
			20					25					30		
Asp	Ala	Phe	Lys	Asn	Lys	Phe	Tyr	Ile	Gly	Val	Ala	Leu	Asn	Arg	Ser
			35				40					45			
Gln	Tyr	Leu	Glu	Gln	Asn	Glu	Gln	Ala	Asp	Lys	Glu	Ile	Lys	Ala	Gln
	50					55					60				
Phe	Ser	Ser	Ile	Val	Ala	Glu	Asn	Cys	Met	Lys	Ser	Glu	Asn	Leu	Glu
65					70				75					80	
Pro	Lys	Glu	Gly	Lys	Phe	Phe	Phe	Asp	Asp	Ala	Asp	Arg	Phe	Val	Ala
				85				90						95	
Phe	Gly	Glu	Lys	Asn	Gly	Met	Tyr	Ile	Ile	Gly	His	Thr	Leu	Ile	Trp
			100					105					110		
His	Ser	Gln	Val	Pro	Lys	Trp	Phe	Phe	Ile	Asp	Asn	Glu	Gly	Lys	Val
	115						120					125			
Val	Ser	Arg	Glu	Val	Leu	Ile	Glu	Arg	Met	Lys	Asn	Tyr	Ile	His	Thr
	130					135					140				
Val	Val	Gly	His	Tyr	Lys	Gly	Arg	Val	Lys	Gly	Trp	Asp	Val	Val	Asn
145					150					155					160
Glu	Ala	Ile	Leu	Asp	Asp	Gly	Ser	Phe	Arg	Gln	Ser	Asn	Phe	Phe	Lys
				165				170						175	
Ile	Leu	Gly	Ala	Asp	Phe	Ile	Lys	Leu	Ala	Phe	Gln	Phe	Ala	His	Glu
			180					185					190		
Ala	Asp	Pro	Asn	Ala	Glu	Leu	Tyr	Asn	Asp	Tyr	Ser	Met	Ser	Asn	
	195						200					205			
Pro	Thr	Lys	Arg	Asp	Gly	Val	Val	Arg	Met	Val	Lys	Ser	Leu	Gln	Gln
	210					215					220				
Gln	Gly	Val	Arg	Ile	Asp	Ala	Ile	Gly	Met	Gln	Gly	His	Val	Gly	Met
225					230					235					240
Asp	Tyr	Pro	Lys	Leu	Asp	Glu	Phe	Glu	Asn	Ser	Ile	Lys	Ala	Phe	Ser
				245					250					255	
Ser	Leu	Gly	Thr	Lys	Val	Met	Ile	Thr	Glu	Leu	Asp	Leu	Ser	Val	Leu
			260					265					270		
Pro	Thr	Pro	Lys	Gly	Lys	Gln	Gly	Ala	Asn	Ile	Ser	Asp	Val	Ala	Ala
	275						280					285			
Tyr	Glu	Glu	Lys	Ile	Asn	Pro	Tyr	Lys	Asn	Gly	Leu	Pro	Ala	Glu	Val
	290					295					300				
Glu	Lys	Ala	Trp	Glu	Asp	Arg	Tyr	Leu	Asp	Phe	Phe	Lys	Leu	Phe	Leu
305					310					315					320
Lys	Tyr	Gln	His	Gln	Ile	Ser	Arg	Val	Thr	Leu	Trp	Gly	Leu	Ser	Asp
				325					330					335	
Gln	Asp	Ser	Trp	Lys	Asn	Asp	Phe	Pro	Val	Arg	Gly	Arg	Thr	Asp	Tyr
			340					345					350		
Pro	Leu	Leu	Phe	Asp	Arg	Gln	Tyr	Lys	Pro	Lys	Pro	Val	Val	Gln	Lys
	355						360					365			

Ile Ile Lys Leu Ala Leu Lys Lys
370 375

<210> 103
<211> 1449
<212> DNA
<213> Bacteria

<220>
<223> Obtained from an environmental sample

<400> 103
atgctgtcac attcccttcc cccgtccacc gtccgccgga aattggggcg cctcggcgcg 60
gcgctgctcg tcggcgccgt cggcgccgcc accgtgctcg tggcgcccct cacctcgcac 120
gccgcccaga gcacgctcgg cgccgcggcg aagcagagcg gccgggtactt cggcaccgcc 180
atcgccctcg gcaggctcaa cgactcgacg tacacgacga tcgcaaccg cgagttcaac 240
tcggtgaccg ccgagaacga gatgaagatc gacgccaccg aaccccagca gggccgcttc 300
gacttcaccg ccggcgaccg cgtctacaac tggcggtgc agaacggcaa gcaggtacgg 360
ggccacaccc tggcctggca ctcccagcag cccgcctgga tgcagaacct cagcggcagc 420
gcgctgcgca cggcgatgac caaccacatc aacggcgta tggcccacta caagggcaag 480
atcgggccagt gggacgtcgt caacgaggcg ttccgggacg gcagttcggg agcgcgccgg 540
gactccaacc tccagcggag cggcaacgac tggatcgagg tcgccttcgg caccgcccgc 600
gccgcccgacc cggccgcca gctctgttac aacgactaca acgtcgagaa ctggacgtgg 660
gccaagaccc aggccatgta cgccatgggtc aaggacttca agcagcgcg cggtgccatc 720
gactgcgtcg gcttccagtc gcacttcaac aacgacagcc cctacaacag caacttcgcg 780
accaccctcc agagtttcgc cgccctcggc gtcgacgtgg ccatcaccga actcgacatc 840
cagggcgccct cgggcacgac ctacgccaac gtgaccaacg actgcctggc cgtcccgcgc 900
tgcttcggca tcaccgtctg ggggtgtccgc gacaccgact cctggcgagc cgagcacact 960
ccgctgtctt tcaacggcga cggcagcaag aagcccgcct actcctccgt cctcaacgcc 1020
ctcaactccg tctcccccaa ccccaacccc actccgaccc cctccccgg cgccgggccc 1080
atcaagggag tcgcctcggg ccgctgcgtg gacgtacccg gagccggcac cgccgacggc 1140
acccaggtcc agctgtggga ctgcaacaac cgcaccaacc agcagtgga cctcaccgcc 1200
gccggtgagc tcagggctta cggcgacaag tgcttgagc cgcggcgac cggcaacggc 1260
gccaaggtcc agatctacag ctgctggggc ggcgacaacc agaagtggcg cctcaactcc 1320
gacggttcca tcgtcgggtg ccagtcggcg cttctgctcg acgcccgtgc cggcggcacc 1380
gccaacggca cgctgatcca gctctactcc tgctggaaca gcggcaacca gcgctggacc 1440
cgcacctga 1449

<210> 104
<211> 482
<212> PRT
<213> Bacteria

<220>
<223> Obtained from an environmental sample

<221> SIGNAL
<222> (1)...(41)

<400> 104
Met Arg Ser His Ser Leu Pro Pro Ser Thr Val Arg Arg Lys Leu Gly
1 5 10 15
Gly Leu Gly Ala Ala Leu Leu Val Gly Ala Val Gly Ala Ala Thr Val
20 25 30
Leu Val Ala Pro Leu Thr Ser His Ala Ala Glu Ser Thr Leu Gly Ala
35 40 45
Ala Ala Lys Gln Ser Gly Arg Tyr Phe Gly Thr Ala Ile Ala Ser Gly
50 55 60
Arg Leu Asn Asp Ser Thr Tyr Thr Thr Ile Ala Asn Arg Glu Phe Asn
65 70 75 80
Ser Val Thr Ala Glu Asn Glu Met Lys Ile Asp Ala Thr Glu Pro Gln
85 90 95
Gln Gly Arg Phe Asp Phe Thr Ala Gly Asp Arg Val Tyr Asn Trp Ala
100 105 110
Val Gln Asn Gly Lys Gln Val Arg Gly His Thr Leu Ala Trp His Ser
115 120 125
Gln Gln Pro Ala Trp Met Gln Asn Leu Ser Gly Ser Ala Leu Arg Thr
130 135 140
Ala Met Thr Asn His Ile Asn Gly Val Met Ala His Tyr Lys Gly Lys

145	Ile	Gly	Gln	Trp	Asp	150	Val	Val	Asn	Glu	Ala	155	Phe	Ala	Asp	Gly	Ser	160	Ser
					165							170						175	
	Gly	Ala	Arg	Arg	Asp	Ser	Asn	Leu	Gln	Arg	Ser	Gly	Asn	Asp	Trp	Ile			
					180														
	Glu	Val	Ala	Phe	Arg	Thr	Ala	Arg	200	Ala	Ala	Asp	Pro	Ala	Lys	Leu			
					195														
	Cys	Tyr	Asn	Asp	Tyr	Asn	Val	Glu	215	Asn	Trp	Thr	Trp	Ala	Lys	Thr	Gln		
					210														
	Ala	Met	Tyr	Ala	Met	Val	Lys	Asp	Phe	Lys	Gln	Arg	Gly	Val	Pro	Ile			
					225														
	Asp	Cys	Val	Gly	Phe	Gln	Ser	His	Phe	Asn	Asn	Asp	Ser	Pro	Tyr	Asn			
					245														
	Ser	Asn	Phe	Arg	Thr	Thr	Leu	Gln	Ser	Phe	Ala	Ala	Leu	Gly	Val	Asp			
					260														
	Val	Ala	Ile	Thr	Glu	Leu	Asp	Ile	280	Gln	Gly	Ala	Ser	Gly	Thr	Thr	Tyr		
					275														
	Ala	Asn	Val	Thr	Asn	Asp	Cys	Leu	295	Ala	Val	Pro	Arg	Cys	Leu	Gly	Ile		
					290														
	Thr	Val	Trp	Gly	Val	Arg	Asp	Thr	310	Asp	Ser	Trp	Arg	Ala	Glu	His	Thr		
					305														
	Pro	Leu	Leu	Phe	Asn	Gly	Asp	Gly	325	Ser	Lys	Lys	Pro	Ala	Tyr	Ser	Ser		
					320														
	Val	Leu	Asn	Ala	Leu	Asn	Ser	Val	340	Ser	Pro	Asn	Pro	Asn	Pro	Thr	Pro		
					335														
	Thr	Pro	Ser	Pro	Gly	Ala	Gly	Pro	360	Ile	Lys	Gly	Val	Ala	Ser	Gly	Arg		
					355														
	Cys	Val	Asp	Val	Pro	Gly	Ala	Gly	375	Thr	Ala	Asp	Gly	Thr	Gln	Val	Gln		
					370														
	Leu	Trp	Asp	Cys	Asn	Asn	Arg	Thr	390	Asn	Gln	Gln	Trp	Thr	Leu	Thr	Ala		
					385														
	Ala	Gly	Glu	Leu	Arg	Val	Tyr	Gly	405	Asp	Lys	Cys	Leu	Asp	Ala	Ala	Gly		
					400														
	Thr	Gly	Asn	Gly	Ala	Lys	Val	Gln	420	Ile	Tyr	Ser	Cys	Trp	Gly	Gly	Asp		
					415														
	Asn	Gln	Lys	Trp	Arg	Leu	Asn	Ser	440	Asp	Gly	Ser	Ile	Val	Gly	Val	Gln		
					435														
	Ser	Gly	Leu	Cys	Leu	Asp	Ala	Ala	455	Gly	Gly	Thr	Ala	Asn	Gly	Thr			
					450														
	Leu	Ile	Gln	Leu	Tyr	Ser	Cys	Trp	470	Asn	Ser	Gly	Asn	Gln	Arg	Trp	Thr		
					465														
	Arg	Thr							475										

<210> 105
 <211> 2793
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 105																			
atgaagttca	ctttgatgcc	gctgctgtgc	gggttcgcct	tgctgttggg	ttgcgcggtg														60
caggcaacc	cagccgcttc	gttacagcag	gcttatcagc	cgtatttcca	tatcggtact														120
gccgtcagct	tggcgcaact	gcaagcatcg	aaaaaccatg	aacgagattt	aatcgcccag														180
cactttaaca	gtctgaccgc	tgaaaacctg	atgaaatggg	aaaaaatcca	accgactgaa														240
ggcaactttg	attttacagc	ggccgacaag	ctcgtcgctt	ttgctgaaca	acatcggtatg														300
tggctggctg	gccatacgat	cctgtggcat	gaacaaaccc	cggactgggt	atttcagggg														360
ccagatggca	aaccggccag	caagcaagtg	ttactcggca	gattaaaaaa	gcatatccaa														420
actgtggctg	gtcgtttacca	aggtcgggta	catggctggg	atgtagtgaa	tgaagcgctg														480
aatgaagatg	gcagtctgcg	cgatacgccg	tggcgaaaaa	ttctgggtga	tgattacatt														540
gccaccactt	ttgcgctggg	gcatcaggct	gaccccaaag	ccaaactcta	ttacaacgac														600
tacaacctgt	ataaaccaaa	aaaacgcact	ggcgtgctac	ggatcatcca	gcaactgcag														660
caacaacaag	tgccatttca	tgccattggc	gaacaagcgc	attatggtct	cgattcgccg														720
aaattgcagg	aagttgaaga	ctcgatcaac	gcctttgcag	ccaccggcct	cgacgtgatg														780
ctgaccgagt	tggaaatttc	ggtgctaccg	tttcgcgctg	gcatgacacc	aggcgccgat														840
atcagtcagc	atcaggaact	gcaacaacag	ctgaatcctt	accgcgaagg	cttaccaaaa														900
accgtcgaac	aggcctggca	acaacgttat	ctggatctgt	tttcgctggt	attgcgccag														960

catcaaaaat	tacaccgggt	gacgttttgg	ggtttagatg	atggccaaag	ctggcgcaat	1020
aacttttcaa	tgcgcggtcg	taccgattac	ccgctactgt	ttgaccgcaa	gctgcaagcc	1080
aaaccgctat	tgagcgctat	gatcaaaactg	gcagaaactc	aagcctcagc	caagccgaaa	1140
gtaaatacagc	tcggttttgc	gccaaatgcg	caaaaattgc	tggtggtgcc	ggggcggcag	1200
gcggtgtcgt	ttcagatcat	caatcaaagc	aacggcaaaa	cggtgttgca	aggccaaagt	1260
tcggtggctc	agttttggcc	cgaatcgggc	gagtggttca	gtatcgctga	cttttcgacc	1320
ttaaccaccc	aagggcggtta	tcaggttgaa	gcggctggat	taactccgat	caccgtcgag	1380
attactgctg	aaccttatgc	cgcgctgcat	gatgctgcca	tcaaagccta	ttattttaat	1440
cgcgctcgc	tggcgctgga	gccaagtittt	gccgggctt	gggcgcgcgc	tgccggtcat	1500
ccggataaca	aagtgttggt	gcacacttcc	gccgcttccg	acaagcgacc	agccggtttt	1560
gtgatcagcg	ccgctaaagg	ctggtatgac	gccggtgact	ataacaaata	cggtgtcaat	1620
tccggtattt	ccagttacac	cctattgcaa	gcctggcagg	attttcctga	gttttatcgc	1680
gacagaacat	ggaatcttcc	ggagtccagc	aacaacctac	cagacattct	cgacgagacg	1740
ttatggaatt	tacagtggct	gagcaccatg	caggacccaa	gcgacggcgg	cggtgtatcac	1800
aagctgactg	aactgaattt	ctctgctacc	caaatgccgt	cagaagtgac	agcgccacgt	1860
tatgtgggtg	aaaaaaccac	ggcagcggca	cttaatttcg	cggcggtgct	ggccaaagcc	1920
agtcgcattt	ttacagaatt	tgaaacgcaa	ctgcccgcc	tgtcacagca	atatcgccag	1980
caagcattag	cagcctggca	atgggcgcaa	aaaaatccac	aacaaattta	tcagcaacca	2040
gccgatgttc	acaccggtgc	ttatggcgac	aaacagctgg	ctgatgaatg	ggcttgggct	2100
ggggcggagc	tatatattt	aaccggtgag	cagagttacc	tgcagccgtt	gttggcactt	2160
gagacgccaa	tcaccgcagc	ttcctgggcc	aatgtggcgg	cgttgggtta	ttttgcgttg	2220
gcatccgctg	aacagtttga	gcctgcactt	cgaaaaaaag	tgcagcaaaa	aatccaacaa	2280
gccgcggcgc	aaattgtagc	cgagcatcaa	gcgtccgcct	accaggtggc	gatgactcaa	2340
aaagattttg	tctggggcag	taatgcggtg	gcgatgaaca	aaggcatgtt	gttatatcaa	2400
gcgtggaaaa	ttgaccaca	accagagctg	cgacaggcga	tgcaagggct	gctggattac	2460
gtcctcggtc	gcaaccgcct	gcagctgtct	tatgtcacag	gttttggtgc	gcaaagcccg	2520
caacatatcc	atcaccgccc	ctcggcgcca	gatcagatca	aagcaccagt	gccgggctgg	2580
ttagtgggtg	gtgcacagcc	gggtaagcaa	gataaatgct	cttattccgg	tatttttgct	2640
accggcactt	taccgctgc	cagcacttta	cctgcaacga	cttatctcga	ccactggtgc	2700
agctacgcca	ccaatgaagt	ggcgattaac	tggaatgcac	ctttagtgtg	cgtgctggcc	2760
tggagccttt	caccagactc	catgaccaa	tga			2793

<210> 106

<211> 930

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(22)

<400> 106

Met	Lys	Phe	Thr	Leu	Met	Pro	Leu	Leu	Cys	Gly	Phe	Ala	Leu	Leu	Leu
1				5				10						15	
Gly	Cys	Ala	Val	Gln	Ala	Thr	Pro	Ala	Ala	Ser	Leu	Gln	Gln	Ala	Tyr
			20					25					30		
Gln	Pro	Tyr	Phe	His	Ile	Gly	Thr	Ala	Val	Ser	Leu	Ala	Gln	Leu	Gln
			35				40					45			
Ala	Ser	Lys	Asn	His	Glu	Arg	Asp	Leu	Ile	Ala	Gln	His	Phe	Asn	Ser
			50			55					60				
Leu	Thr	Ala	Glu	Asn	Leu	Met	Lys	Trp	Glu	Lys	Ile	Gln	Pro	Thr	Glu
65				70					75					80	
Gly	Asn	Phe	Asp	Phe	Thr	Ala	Ala	Asp	Lys	Leu	Val	Ala	Phe	Ala	Glu
			85					90					95		
Gln	His	Arg	Met	Trp	Leu	Val	Gly	His	Thr	Ile	Leu	Trp	His	Glu	Gln
			100					105					110		
Thr	Pro	Asp	Trp	Val	Phe	Gln	Gly	Pro	Asp	Gly	Lys	Pro	Ala	Ser	Lys
			115				120					125			
Gln	Val	Leu	Leu	Gly	Arg	Leu	Lys	Lys	His	Ile	Gln	Thr	Val	Val	Gly
			130			135				140					
Arg	Tyr	Gln	Gly	Arg	Val	His	Gly	Trp	Asp	Val	Val	Asn	Glu	Ala	Leu
145				150					155					160	
Asn	Glu	Asp	Gly	Ser	Leu	Arg	Asp	Thr	Pro	Trp	Arg	Lys	Ile	Leu	Gly
			165					170					175		
Asp	Asp	Tyr	Ile	Ala	Thr	Thr	Phe	Ala	Leu	Val	His	Gln	Val	Asp	Pro
			180					185					190		
Lys	Ala	Lys	Leu	Tyr	Tyr	Asn	Asp	Tyr	Asn	Leu	Tyr	Lys	Pro	Lys	Lys

Arg	Thr	195	Gly	Val	Leu	Arg	Ile	200	Gln	Gln	Leu	Gln	205	Gln	Gln	Gln	Val
210	Ile	His	Ala	Ile	Gly	Glu	Gln	Ala	His	Tyr	Gly	Leu	Asp	Ser	Pro		
225	Leu	Gln	Glu	Val	230	Glu	Asp	Ser	Ile	Asn	Ala	Phe	Ala	Ala	Thr	240	Gly
	Leu	Asp	Val	Met	245	Leu	Thr	Glu	Leu	Glu	Ile	Ser	Val	Leu	Pro	Phe	Pro
	Pro	Gly	Met	260	Thr	Pro	Gly	Ala	Asp	Ile	Ser	Gln	His	Gln	Glu	Leu	Gln
	Gln	Gln	275	Leu	Asn	Pro	Tyr	Arg	Glu	Gly	Leu	Pro	Lys	Thr	Val	Glu	Gln
	290	Ala	Trp	Gln	Gln	Arg	Tyr	Leu	Asp	Leu	Phe	Ser	Leu	Leu	Leu	Arg	Gln
305	His	Gln	Lys	Leu	His	310	Arg	Val	Thr	Phe	Trp	Gly	Leu	Asp	Asp	Gly	Gln
	Ser	Trp	Arg	Asn	Asn	Phe	Pro	Met	Arg	Gly	Arg	Thr	Asp	Tyr	Pro	Leu	
	Leu	Phe	Asp	Arg	Lys	Leu	Gln	Ala	Lys	Pro	Leu	Leu	Ser	Ala	Leu	Ile	
	Lys	Leu	Ala	Glu	Thr	Gln	Ala	Ser	Ala	Lys	Pro	Lys	Val	Asn	Gln	Leu	
	Gly	Phe	Ala	Pro	Asn	Ala	Gln	Lys	Leu	Leu	Val	Val	Pro	Gly	Arg	Gln	
385	Ala	Val	Ser	Phe	Gln	405	Ile	Ile	Asn	Gln	Ser	Asn	Gly	Lys	Thr	Val	Leu
	Gln	Gly	Gln	Ser	Ser	Val	Ala	Gln	Phe	Trp	Pro	Glu	Ser	Gly	Glu	Trp	
	Val	Ser	Ile	Ala	Asp	Phe	Ser	Thr	Leu	Thr	Thr	Gln	Gly	Arg	Tyr	Gln	
	Val	Glu	Ala	Ala	Gly	Leu	Thr	Pro	Ile	Thr	Val	Glu	Ile	Thr	Ala	Glu	
	Pro	Tyr	Ala	Ala	Leu	His	Asp	Ala	Ser	Ile	Lys	Ala	Tyr	Tyr	Phe	Asn	
465	Arg	Ala	Ser	Leu	Ala	Leu	Glu	Pro	Ser	Phe	Ala	Gly	Pro	Trp	Ala	Arg	
	Ala	Ala	Gly	His	Pro	Asp	Asn	Lys	Val	Leu	Val	His	Thr	Ser	Ala	Ala	
	Ser	Asp	Lys	Arg	Pro	Ala	Gly	Phe	Val	Ile	Ser	Ala	Ala	Lys	Gly	Trp	
	Tyr	Asp	Ala	Gly	Asp	Tyr	Asn	Lys	Tyr	Val	Val	Asn	Ser	Gly	Ile	Ser	
	Ser	Tyr	Thr	Leu	Leu	Gln	Ala	Trp	Gln	Asp	Phe	Pro	Glu	Phe	Tyr	Arg	
545	Asp	Arg	Thr	Trp	Asn	Leu	Pro	Glu	Ser	Ser	Asn	Asn	Leu	Pro	Asp	Ile	
	Leu	Asp	Glu	Thr	Leu	Trp	Asn	Leu	Gln	Trp	Leu	Ser	Thr	Met	Gln	Asp	
	Pro	Ser	Asp	Gly	Gly	Val	Tyr	His	Lys	Leu	Thr	Glu	Leu	Asn	Phe	Ser	
	Ala	Thr	Gln	Met	Pro	Ser	Glu	Val	Thr	Ala	Pro	Arg	Tyr	Val	Val	Gln	
	Lys	Thr	Thr	Ala	Ala	Ala	Leu	Asn	Phe	Ala	Ala	Val	Leu	Ala	Lys	Ala	
625	Ser	Arg	Ile	Phe	Thr	Glu	Phe	Glu	Thr	Gln	Leu	Pro	Gly	Leu	Ser	Gln	
	Gln	Tyr	Arg	Gln	Ala	Leu	Ala	Ala	Trp	Gln	Trp	Ala	Gln	Lys	Asn		
	Pro	Gln	Gln	Ile	Tyr	Gln	Gln	Pro	Ala	Asp	Val	His	Thr	Gly	Ala	Tyr	
	Gly	Asp	Lys	Gln	Leu	Ala	Asp	Glu	Trp	Ala	Trp	Ala	Gly	Ala	Glu	Leu	
	Tyr	Leu	Leu	Thr	Gly	Glu	Gln	Ser	Tyr	Leu	Gln	Pro	Leu	Leu	Ala	Leu	
705	Glu	Thr	Pro	Ile	Thr	Ala	Ala	Ser	Trp	Ala	Asn	Val	Ala	Ala	Leu	Gly	
	Tyr	Phe	Ala	Leu	Ala	Ser	Ala	Glu	Gln	Phe	Glu	Pro	Ala	Leu	Arg	Lys	
				740					745					750			

Lys Val Gln Gln Lys Ile Gln Gln Ala Ala Ala Gln Ile Val Ala Glu
 755 760 765
 His Gln Ala Ser Ala Tyr Gln Val Ala Met Thr Gln Lys Asp Phe Val
 770 775 780
 Trp Gly Ser Asn Ala Val Ala Met Asn Lys Gly Met Leu Leu Tyr Gln
 785 790 795 800
 Ala Trp Lys Ile Asp Pro Gln Pro Glu Leu Arg Gln Ala Met Gln Gly
 805 810 815
 Leu Leu Asp Tyr Val Leu Gly Arg Asn Pro Leu Gln Leu Ser Tyr Val
 820 825 830
 Thr Gly Phe Gly Ala Gln Ser Pro Gln His Ile His His Arg Pro Ser
 835 840 845
 Ala Ala Asp Gln Ile Lys Ala Pro Val Pro Gly Trp Leu Val Gly Gly
 850 855 860
 Ala Gln Pro Gly Lys Gln Asp Lys Cys Ser Tyr Ser Gly Ile Phe Ala
 865 870 875 880
 Thr Gly Thr Leu Pro Ala Ala Ser Thr Leu Pro Ala Thr Thr Tyr Leu
 885 890 895
 Asp His Trp Cys Ser Tyr Ala Thr Asn Glu Val Ala Ile Asn Trp Asn
 900 905 910
 Ala Pro Leu Val Tyr Val Leu Ala Trp Ser Leu Ser Pro Asp Ser Met
 915 920 925
 Thr Lys
 930

<210> 107
 <211> 1725
 <212> DNA
 <213> Bacteria

<400> 107
 gtgtggaagc ccggattgtg gaatttcctt caaatggcag atgaagccgg attgacgagg 60
 gatgaaaca ctccggttcc gacacccagt ccaagccgg ctaacacacg tattgaagcg 120
 gaagattatg acggtattaa ttcttcaagt attgagataa taggtgttcc acctgaagga 180
 ggcagaggaa taggttatat taccagtggg gattatctgg tatacaagag tatagacttt 240
 ggaaacggag caacgtcgtt taaggccaag gttgcaaag caaatacttc caatattgaa 300
 cttagattaa acggtccgaa tggctacttc ataggcacac tctcggtaaa atccacagga 360
 gattggaata catatgagga gcaaacctgc agcattagca aagtaccgg aataaatgat 420
 ttgtacttgg tattcaaagg ccctgtaaac atagactggg tcacttttgg cgttgaaagc 480
 agttccacag gtctggggga tttaaatggg gacggaaata ttaactcgtc ggaccttcag 540
 gcgttaaaga ggcatttgct cggatatatca ccgcttacgg gagaggctct ttaagagcg 600
 gatgtaaata ggagcggcaa agtggattct actgactatt cagtgtgaa aagatatata 660
 ctccgcatta ttacagagtt ccccgacaa ggtgatgtac agacaccaa tccgtctgtt 720
 actccgacac aaactcctat cccacgatt tcgggaaatg ctcttaggga ttatgcggag 780
 gcaaggggaa taataatcgg aacatgtgtc aactatccgt tttacaacaa ttcagatcca 840
 acctacaaca gcattttgca aagagaattt tcaatgggtg tatgtgaaaa tgaaatgaag 900
 tttgatgctt tgcagccgag acaaacggtt tttgattttt cgaaaggaga ccagttgctt 960
 gcttttgcag aaagaaacgg tatgcagatg aggggacata cgttgatttg gcacaatcaa 1020
 aaccgcgtcat ggcttacaaa cggttaactgg aaccgggatt cgctgcttgc ggtaatgaaa 1080
 aatcacatta ccactgttat gacccattac aaagggtaaa ttgttgagtg ggatgtggca 1140
 aacgaatgta tggatgattc cggcaacggc ttaagaagca gcatatggag aaatgtaatc 1200
 ggtcaggact accttgacta tgctttcagg tatgcaagag aagcagatcc cgatgcactt 1260
 cttttctaca atgattataa tattgaagac ttgggtccaa agtccaatgc ggtatttaac 1320
 atgattaaaa gtatgaagga aagaggtgtg ccgattgacg gagtaggatt ccaatgccac 1380
 tttatcaatg gaatgagccc cgagtacctt gccagcattg atcaaaatat taagagatat 1440
 gcggaaatag gcgttatagt atcctttacc gaaatagata tacgcatacc tcagtcggaa 1500
 aaccgggcaa ctgcattcca ggtacaggca acaactata aggaacttat gaaaatttgt 1560
 ctggcaaac ccaattgcaa tacctttgta atgtggggat tcacagataa atacacatgg 1620
 attccgggaa ctttcccagg atatggcaat ccattgattt atgacagcaa ttacaatccg 1680
 aaaccggcat acaatgcaat aaaggaagct cttatgggct attga 1725

<210> 108
 <211> 574
 <212> PRT
 <213> Bacteria

<400> 108
 Val Trp Lys Pro Gly Leu Trp Asn Phe Leu Gln Met Ala Asp Glu Ala
 1 5 10 15

Gly Leu Thr Arg Asp Gly Asn Thr Pro Val Pro Thr Pro Ser Pro Lys
 20 25 30
 Pro Ala Asn Thr Arg Ile Glu Ala Glu Asp Tyr Asp Gly Ile Asn Ser
 35 40 45
 Ser Ser Ile Glu Ile Ile Gly Val Pro Pro Glu Gly Gly Arg Gly Ile
 50 55 60
 Gly Tyr Ile Thr Ser Gly Asp Tyr Leu Val Tyr Lys Ser Ile Asp Phe
 65 70 75 80
 Gly Asn Gly Ala Thr Ser Phe Lys Ala Lys Val Ala Asn Ala Asn Thr
 85 90 95
 Ser Asn Ile Glu Leu Arg Leu Asn Gly Pro Asn Gly Thr Leu Ile Gly
 100 105 110
 Thr Leu Ser Val Lys Ser Thr Gly Asp Trp Asn Thr Tyr Glu Glu Gln
 115 120 125
 Thr Cys Ser Ile Ser Lys Val Thr Gly Ile Asn Asp Leu Tyr Leu Val
 130 135 140
 Phe Lys Gly Pro Val Asn Ile Asp Trp Phe Thr Phe Gly Val Glu Ser
 145 150 155 160
 Ser Ser Thr Gly Leu Gly Asp Leu Asn Gly Asp Gly Asn Ile Asn Ser
 165 170 175
 Ser Asp Leu Gln Ala Leu Lys Arg His Leu Leu Gly Ile Ser Pro Leu
 180 185 190
 Thr Gly Glu Ala Leu Leu Arg Ala Asp Val Asn Arg Ser Gly Lys Val
 195 200 205
 Asp Ser Thr Asp Tyr Ser Val Leu Lys Arg Tyr Ile Leu Arg Ile Ile
 210 215 220
 Thr Glu Phe Pro Gly Gln Gly Asp Val Gln Thr Pro Asn Pro Ser Val
 225 230 235 240
 Thr Pro Thr Gln Thr Pro Ile Pro Thr Ile Ser Gly Asn Ala Leu Arg
 245 250 255
 Asp Tyr Ala Glu Ala Arg Gly Ile Lys Ile Gly Thr Cys Val Asn Tyr
 260 265 270
 Pro Phe Tyr Asn Asn Ser Asp Pro Thr Tyr Asn Ser Ile Leu Gln Arg
 275 280 285
 Glu Phe Ser Met Val Val Cys Glu Asn Glu Met Lys Phe Asp Ala Leu
 290 295 300
 Gln Pro Arg Gln Asn Val Phe Asp Phe Ser Lys Gly Asp Gln Leu Leu
 305 310 315 320
 Ala Phe Ala Glu Arg Asn Gly Met Gln Met Arg Gly His Thr Leu Ile
 325 330 335
 Trp His Asn Gln Asn Pro Ser Trp Leu Thr Asn Gly Asn Trp Asn Arg
 340 345 350
 Asp Ser Leu Leu Ala Val Met Lys Asn His Ile Thr Thr Val Met Thr
 355 360 365
 His Tyr Lys Gly Lys Ile Val Glu Trp Asp Val Ala Asn Glu Cys Met
 370 375 380
 Asp Asp Ser Gly Asn Gly Leu Arg Ser Ser Ile Trp Arg Asn Val Ile
 385 390 395 400
 Gly Gln Asp Tyr Leu Asp Tyr Ala Phe Arg Tyr Ala Arg Glu Ala Asp
 405 410 415
 Pro Asp Ala Leu Leu Phe Tyr Asn Asp Tyr Asn Ile Glu Asp Leu Gly
 420 425 430
 Pro Lys Ser Asn Ala Val Phe Asn Met Ile Lys Ser Met Lys Glu Arg
 435 440 445
 Gly Val Pro Ile Asp Gly Val Gly Phe Gln Cys His Phe Ile Asn Gly
 450 455 460
 Met Ser Pro Glu Tyr Leu Ala Ser Ile Asp Gln Asn Ile Lys Arg Tyr
 465 470 475 480
 Ala Glu Ile Gly Val Ile Val Ser Phe Thr Glu Ile Asp Ile Arg Ile
 485 490 495
 Pro Gln Ser Glu Asn Pro Ala Thr Ala Phe Gln Val Gln Ala Asn Asn
 500 505 510
 Tyr Lys Glu Leu Met Lys Ile Cys Leu Ala Asn Pro Asn Cys Asn Thr
 515 520 525
 Phe Val Met Trp Gly Phe Thr Asp Lys Tyr Thr Trp Ile Pro Gly Thr
 530 535 540
 Phe Pro Gly Tyr Gly Asn Pro Leu Ile Tyr Asp Ser Asn Tyr Asn Pro
 545 550 555 560
 Lys pro Ala Tyr Asn Ala Ile Lys Glu Ala Leu Met Gly Tyr

565

570

<210> 109
 <211> 1242
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 109
 atgctaaaag ttttacgtaa acctattatt tctggattag ctttagctct attattgccg 60
 gcaggggcag ctgggtgccg aactaatatt tcaaagaagc caaatataag tggattaacc 120
 gcgcccgaat tagaccaaag atataaagat tctttcacca ttgggtgctgc ggttgagccg 180
 tatcaattat tagatgcaaa agattcaca atgctaaagc ggcattttta tagtatcgta 240
 gcagagaatg tcatgaagcc tagtagttta cagccagtag aaggacaatt caattgggag 300
 ccggccgata aacttggtca gtttgcgaag gaaaatggaa tggacatgcg cggacatacg 360
 cttgtctggc aatctccagg accggattgg ttctttgaag atgcggcagg aaatccaatg 420
 gttgtttggg aaaatggcag gcaagtgggt gccgatccag caaatcttca ggaaaacaaa 480
 gagctcttac ttagccgatt acaaaatcat attcaggcag tcgtaacgcg ttataaagat 540
 gatataaaat cttgggatgt tgttaatgaa gtaatcgatg aatggggcgg acattctgaa 600
 gggctgctgc aatctccatg gttcctcatc accggaacgg actatattaa agttgctttt 660
 gaaactgcaa gagaatatgc agctccagac gctaagctgt atatcaatga ttacaatata 720
 gaagtagaac caaaaaggac gcacctttat aacttagtaa aaagttaaaga agaagaacaa 780
 aacgttccaa ttgatgggtg tgggcatcag tctcacattc aaattggctg gccttcagaa 840
 aaagaaattg aagataccat taatatgttt gcagatcttg gtttagataa ccaaatcacc 900
 gagcttgatg ttagtatgta tggctggcca gtaaggctgt atccaactta tgatgcatc 960
 ccagaactta aattcatgga tcaagcagct cgttatgatc gtttatttaa gttatatgag 1020
 aaattaggag ataaaatcag taatgtgaca ttctggggta ttgcggataa ccatacatgg 1080
 ctgaatgacc gtgcagatgt ttactatgat gaaaatggaa atgttgatt agatagagaa 1140
 acaccaagag tagaaagagg agcaggaaaa gatgcgcatc ttgtatttga tcctgaatac 1200
 aatgtaaaac cagcttattg ggcaattatc gaccacaaat aa 1242

<210> 110
 <211> 413
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(26)

<400> 110
 Met Leu Lys Val Leu Arg Lys Pro Ile Ile Ser Gly Leu Ala Leu Ala
 1 5 10 15
 Leu Leu Leu Pro Ala Gly Ala Ala Gly Ala Glu Thr Asn Ile Ser Lys
 20 25 30
 Lys Pro Asn Ile Ser Gly Leu Thr Ala Pro Gln Leu Asp Gln Arg Tyr
 35 40 45
 Lys Asp Ser Phe Thr Ile Gly Ala Ala Val Glu Pro Tyr Gln Leu Leu
 50 55 60
 Asp Ala Lys Asp Ser Gln Met Leu Lys Arg His Phe Asn Ser Ile Val
 65 70 75 80
 Ala Glu Asn Val Met Lys Pro Ser Ser Leu Gln Pro Val Glu Gly Gln
 85 90 95
 Phe Asn Trp Glu Pro Ala Asp Lys Leu Val Gln Phe Ala Lys Glu Asn
 100 105 110
 Gly Met Asp Met Arg Gly His Thr Leu Val Trp His Ser Gln Val Pro
 115 120 125
 Asp Trp Phe Phe Glu Asp Ala Ala Gly Asn Pro Met Val Val Trp Glu
 130 135 140
 Asn Gly Arg Gln Val Val Ala Asp Pro Ala Asn Leu Gln Glu Asn Lys
 145 150 155 160
 Glu Leu Leu Leu Ser Arg Leu Gln Asn His Ile Gln Ala Val Val Thr
 165 170 175
 Arg Tyr Lys Asp Asp Ile Lys Ser Trp Asp Val Val Asn Glu Val Ile
 180 185 190

Asp Glu Trp Gly Gly His Ser Glu Gly Leu Arg Gln Ser Pro Trp Phe
 195 200 205
 Leu Ile Thr Gly Thr Asp Tyr Ile Lys Val Ala Phe Glu Thr Ala Arg
 210 215 220
 Glu Tyr Ala Ala Pro Asp Ala Lys Leu Tyr Ile Asn Asp Tyr Asn Thr
 225 230 235 240
 Glu Val Glu Pro Lys Arg Thr His Leu Tyr Asn Leu Val Lys Ser Leu
 245 250 255
 Lys Glu Glu Gln Asn Val Pro Ile Asp Gly Val Gly His Gln Ser His
 260 265 270
 Ile Gln Ile Gly Trp Pro Ser Glu Lys Glu Ile Glu Asp Thr Ile Asn
 275 280 285
 Met Phe Ala Asp Leu Gly Leu Asp Asn Gln Ile Thr Glu Leu Asp Val
 290 295 300
 Ser Met Tyr Gly Trp Pro Val Arg Ser Tyr Pro Thr Tyr Asp Ala Ile
 305 310 315 320
 Pro Glu Leu Lys Phe Met Asp Gln Ala Ala Arg Tyr Asp Arg Leu Phe
 325 330 335
 Lys Leu Tyr Glu Lys Leu Gly Asp Lys Ile Ser Asn Val Thr Phe Trp
 340 345 350
 Gly Ile Ala Asp Asn His Thr Trp Leu Asn Asp Arg Ala Asp Val Tyr
 355 360 365
 Tyr Asp Glu Asn Gly Asn Val Val Leu Asp Arg Glu Thr Pro Arg Val
 370 375 380
 Glu Arg Gly Ala Gly Lys Asp Ala Pro Phe Val Phe Asp Pro Glu Tyr
 385 390 395 400
 Asn Val Lys Pro Ala Tyr Trp Ala Ile Ile Asp His Lys
 405 410

<210> 111
 <211> 1089
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 111
 atgttgacga ccccgacaac tcaagatcat gtccccgtgc ttaaggacgc tttcaaaggc 60
 aagttcctca ttggagccgt gctgggttat gacgcactcc agggaaagga tccggcgagt 120
 gtggaaattg cgaccacgca cttcgatgct ctactgctgcg aaaacagcat gaagcccgct 180
 ctggtgcaac ctaaagaggg cgaatttgac ttgctgatg gagaccggt tcttgacatc 240
 acacagcagt gcggtgcgac tgcgattggc cacactttgc tctggcacca acagacaccg 300
 aaatggtttt tctgaggggcc agatgaccag cctactaacc gcgagttggc cctggcacgc 360
 atgagaaagc acatcgccac tcttggtggc cgttacaag gtcgattaa gcaatgggat 420
 gtggtgaatg aggcgattag cgatgcagag ggcgagtact tgagaccaa tagtccatgg 480
 ttcaaggctg ttggagaaga tcacattgct caggctttcc gggcagcgca cgaagccgat 540
 cctgacgcca tcctcatcta taacgattac aacatcgagc aggagtacaa gcgtcccaaa 600
 gcgatacgac tgctgaggtc attacttgag caggacgttc cccttcattg cgtgggcattc 660
 caggggccact ggcgtatgga cactctgaat gtggccgaaa tcgaagaagc tatcaaagaa 720
 tttgctgctc tgggtctcaa ggtcatgatc accgagcttg acatcagcgt gctaccgaca 780
 aagtatcagg gagccgatct ctctaccgcg gaagaattga cgcctgaaat caatccctat 840
 acggagggac taccgagaa cgttgcccgg caacatgccg aatgttaccg ccaagtcttc 900
 aaaatgttcc tgtgccacaa ggatgccatt ggccgtgtca cgctctgggg cgttcatgat 960
 ggcagatcat ggttcaatga ctttcccgtc agagggcgca ccgattatcc tctgcttttc 1020
 gaccggcagg gcaaacccaa gccagcattt tttgccgtct tgaaggctgc gcaagatcag 1080
 ccacaatga 1089

<210> 112
 <211> 362
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 112
 Met Leu Thr Thr Pro Thr Thr Gln Asp His Val Pro Val Leu Lys Asp
 1 5 10 15

Ala Phe Lys Gly Lys Phe Leu Ile Gly Ala Val Leu Gly Tyr Asp Ala
 20 25 30
 Leu Gln Gly Lys Asp Pro Ala Ser Val Glu Ile Ala Thr Thr His Phe
 35 40 45
 Asp Ala Leu Thr Ala Glu Asn Ser Met Lys Pro Ala Leu Val Gln Pro
 50 55 60
 Lys Glu Gly Glu Phe Asp Phe Ala Asp Gly Asp Arg Leu Leu Asp Ile
 65 70 75 80
 Thr Gln Gln Cys Gly Ala Thr Ala Ile Gly His Thr Leu Leu Trp His
 85 90 95
 Gln Gln Thr Pro Lys Trp Phe Phe Glu Gly Pro Asp Asp Gln Pro Thr
 100 105 110
 Asn Arg Glu Leu Ala Leu Ala Arg Met Arg Lys His Ile Ala Thr Leu
 115 120 125
 Val Gly Arg Tyr Lys Gly Arg Ile Lys Gln Trp Asp Val Val Asn Glu
 130 135 140
 Ala Ile Ser Asp Ala Glu Gly Glu Tyr Leu Arg Pro Asn Ser Pro Trp
 145 150 155 160
 Phe Lys Ala Val Gly Glu Asp His Ile Ala Gln Ala Phe Arg Ala Ala
 165 170 175
 His Glu Ala Asp Pro Asp Ala Ile Leu Ile Tyr Asn Asp Tyr Asn Ile
 180 185 190
 Glu Gln Glu Tyr Lys Arg Pro Lys Ala Ile Arg Leu Leu Arg Ser Leu
 195 200 205
 Leu Glu Gln Asp Val Pro Leu His Ala Val Gly Ile Gln Gly His Trp
 210 215 220
 Arg Met Asp Thr Leu Asn Val Ala Glu Ile Glu Glu Ala Ile Lys Glu
 225 230 235 240
 Phe Ala Ala Leu Gly Leu Lys Val Met Ile Thr Glu Leu Asp Ile Ser
 245 250 255
 Val Leu Pro Thr Lys Tyr Gln Gly Ala Asp Leu Ser Thr Arg Glu Glu
 260 265 270
 Leu Thr Pro Glu Ile Asn Pro Tyr Thr Glu Gly Leu Pro Glu Asn Val
 275 280 285
 Ala Arg Gln His Ala Glu Cys Tyr Arg Gln Val Phe Lys Met Phe Leu
 290 295 300
 Cys His Lys Asp Ala Ile Gly Arg Val Thr Leu Trp Gly Val His Asp
 305 310 315 320
 Gly Arg Ser Trp Phe Asn Asp Phe Pro Val Arg Gly Arg Thr Asp Tyr
 325 330 335
 Pro Leu Leu Phe Asp Arg Gln Gly Lys Pro Lys Pro Ala Phe Phe Ala
 340 345 350
 Val Leu Lys Ala Ala Gln Asp Gln Pro Gln
 355 360

<210> 113
 <211> 1155
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 113
 atgttaaaag tattgcgtaa accacttttt tctggattag ctttagcgat agtattacct 60
 accggattat ccagtgccta tgcagctgaa aatcaaccag ttagtgcat agatgcagcg 120
 gttgaacttg atgaaagata tgcagaatca ttcgatattg gtgcagccgt tgagccttct 180
 atgcttcaag gaaaagatgc tgaagtatta aagcgtcatt ataacagcat tgtggccgaa 240
 aatgtaatga aaccgattaa tatacagcct gaagaaggaa agttcacttt taaagaaatg 300
 gataaaatcg ttaagtttgc gaaagaaaat aatatgaagc ttcgtggcca tacccttatt 360
 tggcacagtc aagtaccgga gtggttcttc cttgataaag aaggaaataa gatgggtggat 420
 gaaacggatc caaagcagcg cgaaaaaaat aaaaggcttt tacttaagcg tttagaaacg 480
 catattaaaa cgatcgtcaa gcgctataaa aatgatatta aatgatatta gctcctggga cgtgggtcaac 540
 gaggtagtgg atgataacgg gaaattacgt aattcaccct ggtatcaaat cacagggtact 600
 gattatatca aggttgcttt tgaaacagcg gaccgttatg cagggaagaa cgctaagctt 660
 tatatcaatg actacaacac ggaaatagac cctaaaagag aaaccctcta taatcttgtc 720
 aaggaattag tgaaggagg agtcccagtt gatggagtgg gacatcaagc tcatatccaa 780
 atcggctggc caactatagc ggaaatcgag aaaaccatta atatgtttgc agaccttggc 840
 ctgacaatc aaattacaga actagatgtt agcctttatg ggtggccgcc aaagcctgct 900

taccaactt	atgacgaaat	cccggaagt	gaattcgaac	gtcaagctgt	tcgttacgat	960
caactatttg	atttatacga	gagattggga	gataaaatta	gcagtgtgac	attctggggc	1020
gttgctgaca	accatacatg	gttaaatac	cgtgcagaac	aatataatga	cggggtaggc	1080
gtggacgcac	catttgtttt	cgaataaggat	tataatgtaa	aaccagctta	ttgggctatt	1140
atcgatcgcg	attaa					1155

<210> 114
 <211> 384
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 114

Met	Leu	Lys	Val	Leu	Arg	Lys	Pro	Leu	Phe	Ser	Gly	Leu	Ala	Leu	Ala
1				5					10					15	
Ile	Val	Leu	Pro	Thr	Gly	Leu	Ser	Ser	Ala	Tyr	Ala	Ala	Glu	Asn	Gln
			20					25					30		
Pro	Val	Ser	Ala	Leu	Asp	Ala	Ala	Val	Glu	Leu	Asp	Glu	Arg	Tyr	Ala
		35				40						45			
Glu	Ser	Phe	Asp	Ile	Gly	Ala	Ala	Val	Glu	Pro	Ser	Met	Leu	Gln	Gly
	50				55						60				
Lys	Asp	Ala	Glu	Val	Leu	Lys	Arg	His	Tyr	Asn	Ser	Ile	Val	Ala	Glu
65					70					75				80	
Asn	Val	Met	Lys	Pro	Ile	Asn	Ile	Gln	Pro	Glu	Glu	Gly	Lys	Phe	Thr
			85						90					95	
Phe	Lys	Glu	Met	Asp	Lys	Ile	Val	Lys	Phe	Ala	Lys	Glu	Asn	Asn	Met
			100					105					110		
Lys	Leu	Arg	Gly	His	Thr	Leu	Ile	Trp	His	Ser	Gln	Val	Pro	Glu	Trp
	115					120						125			
Phe	Phe	Leu	Asp	Lys	Glu	Gly	Asn	Lys	Met	Val	Asp	Glu	Thr	Asp	Pro
	130					135					140				
Lys	Gln	Arg	Glu	Lys	Asn	Lys	Arg	Leu	Leu	Leu	Lys	Arg	Leu	Glu	Thr
145					150					155				160	
His	Ile	Lys	Thr	Ile	Val	Lys	Arg	Tyr	Lys	Asn	Asp	Ile	Ser	Ser	Trp
			165						170				175		
Asp	Val	Val	Asn	Glu	Val	Val	Asp	Asp	Asn	Gly	Lys	Leu	Arg	Asn	Ser
			180					185					190		
Pro	Trp	Tyr	Gln	Ile	Thr	Gly	Thr	Asp	Tyr	Ile	Lys	Val	Ala	Phe	Glu
	195					200						205			
Thr	Ala	Asp	Arg	Tyr	Ala	Gly	Lys	Asn	Ala	Lys	Leu	Tyr	Ile	Asn	Asp
	210					215					220				
Tyr	Asn	Thr	Glu	Ile	Asp	Pro	Lys	Arg	Glu	Thr	Leu	Tyr	Asn	Leu	Val
225					230					235				240	
Lys	Glu	Leu	Val	Lys	Glu	Gly	Val	Pro	Val	Asp	Gly	Val	Gly	His	Gln
			245						250					255	
Ala	His	Ile	Gln	Ile	Gly	Trp	Pro	Thr	Ile	Ala	Glu	Ile	Glu	Lys	Thr
			260					265					270		
Ile	Asn	Met	Phe	Ala	Asp	Leu	Gly	Leu	Asp	Asn	Gln	Ile	Thr	Glu	Leu
	275						280					285			
Asp	Val	Ser	Leu	Tyr	Gly	Trp	Pro	Pro	Lys	Pro	Ala	Tyr	Pro	Thr	Tyr
	290					295					300				
Asp	Glu	Ile	Pro	Ala	Ser	Glu	Phe	Glu	Arg	Gln	Ala	Val	Arg	Tyr	Asp
305					310					315				320	
Gln	Leu	Phe	Asp	Leu	Tyr	Glu	Arg	Leu	Gly	Asp	Lys	Ile	Ser	Ser	Val
			325						330				335		
Thr	Phe	Trp	Gly	Val	Ala	Asp	Asn	His	Thr	Trp	Leu	Asn	Asp	Arg	Ala
			340					345					350		
Glu	Gln	Tyr	Asn	Asp	Gly	Val	Gly	Val	Asp	Ala	Pro	Phe	Val	Phe	Asp
	355						360					365			
Lys	Asp	Tyr	Asn	Val	Lys	Pro	Ala	Tyr	Trp	Ala	Ile	Ile	Asp	Arg	Asp
	370					375					380				

<210> 115
 <211> 1362

<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample

```

<400> 115
atgacgaacc gtaaatacgaa cgtgcaccgt tcattgaccg atgatttgct cgatgggtgtc      60
ttcgccgagg caaaagcggg caaagttgag aagtaccgtg ccaccgggat ccttggaacg      120
ctattcggat tctctgtggc gtcctccatc atgttggcgg cttgcagcaa cgcacaagag      180
aatgttccac cagttgtctt atccaccgca cagagcaata tcaccagga gaacgttccg      240
ccgctcaaag atgcgtttaa gggcaagttc ttgattggca ccgcggtgag caatcgcttg      300
ctggagggac aagatccggc cacggaagcc ttggtgcgca ggcacttcga tgctctcacg      360
gcggaataac ccatgaagcc ggatgactg caaccgcgcg aaggccagtt caacttcgtc      420
gccgccgacc gtctgtgtgga aatcgcccag caaagcggcg cgaagtggt cggccacacg      480
ctggtctggc actcccaaac gccaggctgg ttcttccagg gtccgaatgg ccagccagcg      540
agtgcagaag tggccctggc gcggatgcga acacacatca agacggtggt gggacgctac      600
aaagggcgca tcaagcagtg ggatgtgttc aacgaagcga tcaacgacgg ccctggcgtg      660
ctgcggcaaa gtccgtggct gcgtgccatc ggcgaagact acatcgccga agcgttccgc      720
gccgcgcacg aagccgatcc tgacgccatt ctggtctaca acgactacaa catcgaactc      780
aactacaagc gtcccaaggc gctggaactg ctaagaagc tcatcgacca gaaggttccg      840
attcatggtg tgggcattca ggctcactgg cgcatgacc cgccgctggc cgagaccgaa      900
gaagccatca aacagttcgc cgcgctgggc ctgaagggtg tgttcaccga actggacatc      960
ggtgtgctgc cactcagta tcagggggct gacatctcgg cgctgaaac catgacaccc      1020
gaacagcaag cggatgatgaa cccttacact cagggttgc cggctgaagt ggcacagcaa      1080
catgccgagc gctaccgaca ggccttcgag ctgttctctg gccacaagga tgtgattggt      1140
cgcgctcacgc tctggggcac gcatgatggc gaatcctggc tgaacggttt tccggtgcgg      1200
ggccgcaccg actatccctt gctcttcgac cgccggtatc agccaaaacc agccttcttc      1260
gccgtcaggc aggttgacac ggcgcatact gtacaaacga ccggtgcgca aaccaagct      1320
acagcgaaga caattcaaaa agcttctcga gagtacttct ag      1362

```

<210> 116
<211> 453
<212> PRT
<213> Unknown

<220>
<223> Obtained from an environmental sample

```

<400> 116
Met Thr Asn Arg Lys Ser Asn Val His Arg Ser Leu Thr Asp Asp Leu
1      5      10      15
Leu Asp Gly Val Phe Ala Glu Ala Lys Ala Gly Lys Val Glu Lys Tyr
20      25      30
Arg Ala Thr Gly Ile Leu Gly Thr Leu Phe Gly Phe Thr Val Ala Ser
35      40      45
Ser Ile Met Leu Ala Ala Cys Ser Asn Ala Gln Glu Asn Val Pro Pro
50      55      60
Val Ala Ser Ser Thr Ala Gln Ser Asn Ile Thr Gln Glu Asn Val Pro
65      70      75      80
Pro Leu Lys Asp Ala Phe Lys Gly Lys Phe Leu Ile Gly Thr Ala Val
85      90      95
Ser Asn Arg Leu Leu Gly Gly Gln Asp Pro Ala Thr Glu Ala Leu Val
100      105      110
Arg Arg His Phe Asp Ala Leu Thr Ala Glu Asn Ala Met Lys Pro Asp
115      120      125
Ala Leu Gln Pro Arg Glu Gly Gln Phe Asn Phe Val Ala Ala Asp Arg
130      135      140
Leu Val Glu Ile Ala Gln Gln Ser Gly Ala Thr Val Val Gly His Thr
145      150      155      160
Leu Val Trp His Ser Gln Thr Pro Gly Trp Phe Phe Gln Gly Pro Asn
165      170      175
Gly Gln Pro Ala Ser Arg Glu Leu Ala Leu Ala Arg Met Arg Thr His
180      185      190
Ile Lys Thr Val Val Gly Arg Tyr Lys Gly Arg Ile Lys Gln Trp Asp
195      200      205
Val Val Asn Glu Ala Ile Asn Asp Gly Pro Gly Val Leu Arg Gln Ser
210      215      220
Pro Trp Leu Arg Ala Ile Gly Glu Asp Tyr Ile Ala Glu Ala Phe Arg

```


225 Ala Ala His Glu Ala 230 Asp Pro Asp Ala Ile 235 Leu Val Tyr Asn Asp 240 Tyr
 Asn Ile Glu Leu 245 Asn Tyr Lys Arg Pro Lys Ala Leu Glu Leu 255 Leu Lys
 Lys Leu Ile 260 Gln Lys Val Pro 265 Ile His Gly Val Gly Ile Gln Ala
 His Trp Arg Met Thr Pro 280 Leu Ala Glu Thr Glu Glu Ala Ile Lys
 Gln Phe Ala Ala Leu Gly Leu Lys Val Met Phe Thr Glu Leu Asp Ile
 305 Gly Val Leu Pro Thr 310 Gln Tyr Gln Gly Ala Asp Ile Ser Ala Arg Glu
 Thr Met Thr Pro Glu Gln Gln Ala Val Met Asn Pro Tyr Thr Gln Gly
 Leu Pro Ala Glu Val Ala Gln Gln His Ala Glu Arg Tyr Arg Gln Ala
 Phe Glu Leu Phe Leu Arg His Lys Asp Val Ile Gly Arg Val Thr Leu
 370 Trp Gly Thr His Asp Gly Glu Ser Trp Leu Asn 380 Phe Pro Val Arg
 385 Gly Arg Thr Asp Tyr Pro Leu Leu Phe Asp Arg Arg Tyr Gln Pro Lys
 Pro Ala Phe Phe Ala Val Arg Gln Val Ala Gln Ala His Thr Val Gln
 Thr Thr Gly Ala Gln Thr Gln Ala Thr Ala Lys Thr Ile Gln Lys Ala
 435 Ser Arg Glu Tyr Phe 440 445 450

<210> 117
 <211> 1437
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 117
 atgacgaacc gtaaattgaa cgtgcaccgt tcattgagcg atgatttgct cgatggcgcc 60
 ttgcgcgagt caaaagcggg caaagttgag aaataccgtg ccacggggat ccttgaacg 120
 ctattcggat tcaactgtggc gtcctccatc atgttgccgg cttgcagcaa cgacaagag 180
 aatgtccac cagttgcttc atccaccgca caaagcaata tcacccagga gaacgttccg 240
 ccgctcaagg atgcgtttta gggcaagttc ttgattggca ccacgcgcgag caatcgcttg 300
 ctgcagggac aagatccagc cacagaagcc ctggtgcgca ggcacttcga cgccctcacg 360
 gcggaataatg ccatgaagcc tgaatgccatg caaccagag aggggtgagtt caactttgcc 420
 gccgctgacc gcctggtgga aatcgccag caaagcggcg ccacggtggt cggccacacc 480
 ttggtctggc atagccaaac gccaagctgg ttcttcagg gtccagatgg ccaaccggcg 540
 agtcgggaac tggccttggc acggatgcga acgcacatca agactgtggt gggacgctac 600
 aaaggacgca tcaagcaatg ggatgtggtc aacgaagcga tcaacgacgg ccctggagtg 660
 ctgcggccat cgccgtggtt gcgcgccatc ggcgaagact tcatcgccga agcgttccgc 720
 gccgcgcacg aagctgatcc cgacgcgatt ctctgtctaca acgactacaa catcgagctc 780
 aactacaagc gtcccaaggc gctggaacta ctgaagagac tcatcgagca gaaggttccg 840
 attcatggtg tgggcattca ggctcactgg cgcattgacc cgccgctggc cgagatggaa 900
 gagaccatca agcagttttc ggctttgggc ttgaaggtaa tgatcaccga gttggacatt 960
 ggtgtattgc caacacaata ccagggtgac gacatctcgg ctgcgcgagac catgacaccc 1020
 gaacagcaag cggatgatgaa ccttacacg cagggtctgc cggctgaagt ggcgcagcaa 1080
 catgccgagc gttatcgta ggcgtttgag ctgttcacgc gttacaagga tgtgattggt 1140
 cgcggtatccc tgtggggcac gcatgatggc gaatcttggc tgaacgggtt tcccgttcgt 1200
 ggccgcacgg attatcctct actgttcgac cgccggtatc agcctaagcc cgcttcttc 1260
 gcggtgcaaa aggtcgcgca ggcgcagaac gcacaggcag caaccgatca agcaccactt 1320
 gcacaaaacc cagttgcgca gaagaaatct gcaccaaggc aggcgggtca aaatcagacc 1380
 actcaaaagc cagtgtgtaca aaagcaagtc gcggcaagtc gggccgcaga aaagtaa 1437

<210> 118
 <211> 478
 <212> PRT
 <213> Unknown

<220>

<223> obtained from an environmental sample

<400> 118

Met Thr Asn Arg Lys Leu Asn Val His Arg Ser Leu Ser Asp Asp Leu
 1 5 10 15
 Leu Asp Gly Ala Phe Ala Glu Ser Lys Ala Gly Lys Val Glu Lys Tyr
 20 25 30
 Arg Ala Thr Gly Ile Leu Gly Thr Leu Phe Gly Phe Thr Val Ala Ser
 35 40 45
 Ser Ile Met Leu Ala Ala Cys Ser Asn Ala Gln Glu Asn Ala Pro Pro
 50 55 60
 Val Ala Ser Ser Thr Ala Gln Ser Asn Ile Thr Gln Glu Asn Val Pro
 65 70 75 80
 Pro Leu Lys Asp Ala Phe Lys Gly Lys Phe Leu Ile Gly Thr Ile Ala
 85 90 95
 Ser Asn Arg Leu Leu Gln Gly Gln Asp Pro Ala Thr Glu Ala Leu Val
 100 105 110
 Arg Arg His Phe Asp Ala Leu Thr Ala Glu Asn Ala Met Lys Pro Asp
 115 120 125
 Ala Met Gln Pro Arg Glu Gly Glu Phe Asn Phe Ala Ala Asp Arg
 130 135 140
 Leu Val Glu Ile Ala Gln Gln Ser Gly Ala Thr Val Val Gly His Thr
 145 150 155 160
 Leu Val Trp His Ser Gln Thr Pro Ser Trp Phe Phe Gln Gly Pro Asp
 165 170 175
 Gly Gln Pro Ala Ser Arg Glu Leu Ala Leu Ala Arg Met Arg Thr His
 180 185 190
 Ile Lys Thr Val Val Gly Arg Tyr Lys Gly Arg Ile Lys Gln Trp Asp
 195 200 205
 Val Val Asn Glu Ala Ile Asn Asp Gly Pro Gly Val Leu Arg Pro Ser
 210 215 220
 Pro Trp Leu Arg Ala Ile Gly Glu Asp Phe Ile Ala Glu Ala Phe Arg
 225 230 235 240
 Ala Ala His Glu Ala Asp Pro Asp Ala Ile Leu Val Tyr Asn Asp Tyr
 245 250 255
 Asn Ile Glu Leu Asn Tyr Lys Arg Pro Lys Ala Leu Glu Leu Leu Lys
 260 265 270
 Arg Leu Ile Glu Gln Lys Val Pro Ile His Gly Val Gly Ile Gln Ala
 275 280 285
 His Trp Arg Met Thr Pro Pro Leu Ala Glu Met Glu Glu Thr Ile Lys
 290 295 300
 Gln Phe Ser Ala Leu Gly Leu Lys Val Met Ile Thr Glu Leu Asp Ile
 305 310 315 320
 Gly Val Leu Pro Thr Gln Tyr Gln Gly Ala Asp Ile Ser Ala Arg Glu
 325 330 335
 Thr Met Thr Pro Glu Gln Gln Ala Val Met Asn Pro Tyr Thr Gln Gly
 340 345 350
 Leu Pro Ala Glu Val Ala Gln Gln His Ala Glu Arg Tyr Arg Gln Ala
 355 360 365
 Phe Glu Leu Phe Met Arg Tyr Lys Asp Val Ile Gly Arg Val Thr Leu
 370 375 380
 Trp Gly Thr His Asp Gly Glu Ser Trp Leu Asn Gly Phe Pro Val Arg
 385 390 395 400
 Gly Arg Thr Asp Tyr Pro Leu Leu Phe Asp Arg Arg Tyr Gln Pro Lys
 405 410 415
 Pro Ala Phe Phe Ala Val Gln Lys Val Ala Gln Ala Gln Asn Ala Gln
 420 425 430
 Ala Ala Thr Asp Gln Ala Pro Leu Ala Gln Asn Pro Val Ala Gln Lys
 435 440 445
 Lys Ser Ala Pro Arg Gln Ala Ala Gln Asn Gln Thr Gln Lys Pro
 450 455 460
 Val Val Gln Lys Gln Ser Ala Ala Ser Arg Ala Ala Glu Lys
 465 470 475

<210> 119

<211> 2559

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 119

atgaaaaaaa	gattgttagc	gttgatagtg	acattagttt	ttattatctc	attgtttaat	60
cccatattca	ccacaccttt	aacaaatgta	gcaaaggctc	aaagtaacca	aacaaattta	120
aaatttgact	ttgaaaacgg	tactcaagg	tggggagcaa	gagggtgttc	aacaactatt	180
gcaaccgttt	acgagcaagc	ttatgaagga	agttattctt	taaagggttc	aggtagaagt	240
tcaacgtggg	atggagcagt	tgtggatc	acatcaagta	tttcagcaaa	tgtcacctat	300
acagtttctt	tatttggtcg	tcacagcgat	gtaaaaccac	aaagattttc	tgtctatgta	360
tatgtcaaag	ataacacagg	cgaaaaatc	atccaggttg	cagacaaagt	ggttatgcc	420
aacttttgg	agtcagctctt	tggaaggttc	acaatcacaa	catcaaattc	aattcaaaaa	480
gtagaacttc	ttgtatgtgt	tccatctaac	aaatctttag	gattttatct	tgacaatgta	540
gttattactt	cagcacaacc	agcttcctcg	gggtgtgtta	aatcttgac	atttgaaagc	600
ggtagcactg	agggttttgt	tcagagaggt	tcagcttcac	tgacagttgt	cgacgggtga	660
tactatcatt	ctccaacaaa	agcattatat	gtgacaggaa	ggacagctac	atggcagggt	720
gcacagatag	atatgacaag	tttgcttgag	aagggcaagg	attatcagtt	tagcatatgg	780
gtatatcaaa	atagtggag	tgatcagaag	ataaccctta	cgatgcaaa	gaagaatgaa	840
gatggaacta	cgagttatga	ttctataaag	tatcagcaaa	cagttccatc	tggtacatgg	900
acagaagtaa	caggttcata	cacagtcct	cagacagcaa	cacagcttat	attctatggt	960
gaatcaccca	atattaat	tgacttctac	cttgatgact	ttacagcgg	tgacaaaaac	1020
ccacctgttg	taaacccagg	gcttggttaa	tcttgacat	ttgaaagcgg	tagcactgag	1080
ggttttgttc	agagaggttc	agcttcattg	acagttgtcg	acggtgtata	ctatcattct	1140
ccaacaaaag	cattgtatgt	gacaggaagg	acagctacat	ggcaggggtg	acagatagat	1200
atgacaagtt	tgcttgagaa	gggcaaggat	tatcagttta	gcataatggg	atatcaaaat	1260
agtgggaagt	atcagaagat	aacccttacg	atgcaaagga	agaatgaaga	tggaactacg	1320
agttatgatt	ctataaagta	tcagcaaaac	gttccatctg	gtacatggac	agaagtaaca	1380
ggttcataca	cagtgcctca	gacagcaaca	cagcttatat	tctatgttga	atcacccaat	1440
attaattttg	acttctacct	tgatgacttt	acagtaatat	ataaaaatcc	agtgcaggta	1500
ccgattgcag	caaaagaacc	cgaatgggaa	attccgtcac	ttgtcagca	atatagtcaa	1560
tattttctca	taggtgttgc	aataccgtat	aaagtacttc	aaaatcctgt	tgaaagagca	1620
atgggtgttaa	aacacttcaa	cagtataaca	gctgaaaatg	agatgaaacc	tgacgctctg	1680
caaagaacag	aagggaactt	tacattcgat	atagcagacc	agtatgtaaa	cttcgcacag	1740
caaaacggta	ttggaattag	agggcatact	ctggatggc	acagccaagt	acctaattgg	1800
ttcttccagc	acagtgatgg	aacttcactt	gatccaagca	atccagatga	taagcaactt	1860
ttgagagata	gattgaaaaa	tcatattcaa	actgttatgt	caagatacaa	agggaaagtc	1920
tatgcatggg	atgttgtaaa	cgaggcaata	gatgaaagcc	agcctgatgg	atttagaaga	1980
agcgaatgg	acagaatact	tggtccaaca	cctgagacaa	atggatttcc	agaatacatt	2040
gtgcttgctt	tcaggtatgc	aagagaggcg	gatccggatg	caaaactttt	ctacaatgac	2100
tacaacacag	agatatctaa	aaaaagacag	tttatatatg	acatggtaaa	aaagctacat	2160
gatatgggtt	taattgatgg	tggttggttg	caagggcata	taaatgttga	ttctccaaca	2220
gtaaaagata	tagaagatac	aatcaatctt	ttctcaacaa	ttcctggact	tgagatacag	2280
gtaacagagc	ttgacataag	cgtttacaca	agcagcagtc	agcgttatga	tacgcttcct	2340
caggatataa	tgataaaaca	agcaatgaag	tttaaagaac	tatttgaaat	gttaaagaga	2400
catagtata	gagtcacaaa	tgtgacactt	tggggactta	aggatgatta	ttcatggctt	2460
tcaaaggata	gaaataactg	gccattgctt	tttgacagca	actaccaggc	aaaatacagc	2520
tactgggcaa	ttcaaaaagc	ttctcagag	tacttctag			2559

<210> 120

<211> 852

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(33)

<400> 120

Met	Lys	Lys	Arg	Leu	Leu	Ala	Leu	Ile	Val	Thr	Leu	Val	Phe	Ile	Ile
1				5				10					15		
Ser	Leu	Phe	Asn	Pro	Ile	Phe	Thr	Thr	Pro	Leu	Thr	Asn	Val	Ala	Lys
			20					25					30		
Ala	Gln	Ser	Asn	Gln	Thr	Asn	Leu	Lys	Phe	Asp	Phe	Glu	Asn	Gly	Thr
		35					40					45			
Gln	Gly	Trp	Gly	Ala	Arg	Gly	Val	Ser	Thr	Thr	Ile	Ala	Thr	Val	Tyr
	50					55					60				

Glu Gln Ala Tyr Glu Gly Ser Tyr Ser Leu Lys Val Ser Gly Arg Ser
 65 70 75 80
 Ser Thr Trp Asp Gly Ala Val Val Asp Ile Thr Ser Ser Ile Ser Ala
 85 90 95
 Asn Val Thr Tyr Thr Val Ser Leu Phe Val Arg His Ser Asp Val Lys
 100 105 110
 Pro Gln Arg Phe Ser Val Tyr Val Tyr Val Lys Asp Asn Thr Gly Glu
 115 120 125
 Lys Tyr Ile Gln Val Ala Asp Lys Val Val Met Pro Asn Phe Trp Lys
 130 135 140
 Gln Leu Phe Gly Lys Phe Thr Ile Thr Thr Ser Asn Pro Ile Gln Lys
 145 150 155 160
 Val Glu Leu Leu Val Cys Val Pro Ser Asn Lys Ser Leu Gly Phe Tyr
 165 170 175
 Leu Asp Asn Val Val Ile Thr Ser Ala Gln Pro Ala Ser Ser Gly Val
 180 185 190
 Val Lys Ser Cys Thr Phe Glu Ser Gly Ser Thr Glu Gly Phe Val Gln
 195 200 205
 Arg Gly Ser Ala Ser Leu Thr Val Val Asp Gly Val Tyr Tyr His Ser
 210 215 220
 Pro Thr Lys Ala Leu Tyr Val Thr Gly Arg Thr Ala Thr Trp Gln Gly
 225 230 235 240
 Ala Gln Ile Asp Met Thr Ser Leu Leu Glu Lys Gly Lys Asp Tyr Gln
 245 250 255
 Phe Ser Ile Trp Val Tyr Gln Asn Ser Gly Ser Asp Gln Lys Ile Thr
 260 265 270
 Leu Thr Met Gln Arg Lys Asn Glu Asp Gly Thr Thr Ser Tyr Asp Ser
 275 280 285
 Ile Lys Tyr Gln Gln Thr Val Pro Ser Gly Thr Trp Thr Glu Val Thr
 290 295 300
 Gly Ser Tyr Thr Val Pro Gln Thr Ala Thr Gln Leu Ile Phe Tyr Val
 305 310 315 320
 Glu Ser Pro Asn Ile Asn Phe Asp Phe Tyr Leu Asp Asp Phe Thr Ala
 325 330 335
 Val Asp Lys Asn Pro Pro Val Val Asn Pro Gly Leu Val Lys Ser Cys
 340 345 350
 Thr Phe Glu Ser Gly Ser Thr Glu Gly Phe Val Gln Arg Gly Ser Ala
 355 360 365
 Ser Leu Thr Val Val Asp Gly Val Tyr Tyr His Ser Pro Thr Lys Ala
 370 375 380
 Leu Tyr Val Thr Gly Arg Thr Ala Thr Trp Gln Gly Ala Gln Ile Asp
 385 390 395 400
 Met Thr Ser Leu Leu Glu Lys Gly Lys Asp Tyr Gln Phe Ser Ile Trp
 405 410 415
 Val Tyr Gln Asn Ser Gly Ser Asp Gln Lys Ile Thr Leu Thr Met Gln
 420 425 430
 Arg Lys Asn Glu Asp Gly Thr Thr Ser Tyr Asp Ser Ile Lys Tyr Gln
 435 440 445
 Gln Thr Val Pro Ser Gly Thr Trp Thr Glu Val Thr Gly Ser Tyr Thr
 450 455 460
 Val Pro Gln Thr Ala Thr Gln Leu Ile Phe Tyr Val Glu Ser Pro Asn
 465 470 475 480
 Ile Asn Phe Asp Phe Tyr Leu Asp Asp Phe Thr Val Ile Asp Lys Asn
 485 490 495
 Pro Val Thr Val Pro Ile Ala Ala Lys Glu Pro Glu Trp Glu Ile Pro
 500 505 510
 Ser Leu Cys Gln Gln Tyr Ser Gln Tyr Phe Ser Ile Gly Val Ala Ile
 515 520 525
 Pro Tyr Lys Val Leu Gln Asn Pro Val Glu Arg Ala Met Val Leu Lys
 530 535 540
 His Phe Asn Ser Ile Thr Ala Glu Asn Glu Met Lys Pro Asp Ala Leu
 545 550 555 560
 Gln Arg Thr Glu Gly Asn Phe Thr Phe Asp Ile Ala Asp Gln Tyr Val
 565 570 575
 Asn Phe Ala Gln Gln Asn Gly Ile Gly Ile Arg Gly His Thr Leu Val
 580 585 590
 Trp His Ser Gln Val Pro Asn Trp Phe Phe Gln His Ser Asp Gly Thr
 595 600 605
 Ser Leu Asp Pro Ser Asn Pro Asp Asp Lys Gln Leu Leu Arg Asp Arg

610 615 620
 Leu Lys Asn His Ile Gln Thr Val Met Ser Arg Tyr Lys Gly Lys Val
 625 630 635 640
 Tyr Ala Trp Asp Val Val Asn Glu Ala Ile Asp Glu Ser Gln Pro Asp
 645 650 655
 Gly Phe Arg Arg Ser Glu Trp Tyr Arg Ile Leu Gly Pro Thr Pro Glu
 660 665 670
 Thr Asn Gly Ile Pro Glu Tyr Ile Val Leu Ala Phe Arg Tyr Ala Arg
 675 680 685
 Glu Ala Asp Pro Asp Ala Lys Leu Phe Tyr Asn Asp Tyr Asn Thr Glu
 690 695 700
 Ile Ser Lys Lys Arg Gln Phe Ile Tyr Asp Met Val Lys Lys Leu His
 705 710 715 720
 Asp Met Gly Leu Ile Asp Gly Val Gly Leu Gln Gly His Ile Asn Val
 725 730 735
 Asp Ser Pro Thr Val Lys Asp Ile Glu Asp Thr Ile Asn Leu Phe Ser
 740 745 750
 Thr Ile Pro Gly Leu Glu Ile Gln Val Thr Glu Leu Asp Ile Ser Val
 755 760 765
 Tyr Thr Ser Ser Ser Gln Arg Tyr Asp Thr Leu Pro Gln Asp Ile Met
 770 775 780
 Ile Lys Gln Ala Met Lys Phe Lys Glu Leu Phe Glu Met Leu Lys Arg
 785 790 795 800
 His Ser Asp Arg Val Thr Asn Val Thr Leu Trp Gly Leu Lys Asp Asp
 805 810 815
 Tyr Ser Trp Leu Ser Lys Asp Arg Asn Asn Trp Pro Leu Leu Phe Asp
 820 825 830
 Ser Asn Tyr Gln Ala Lys Tyr Ser Tyr Trp Ala Ile Gln Lys Ala Ser
 835 840 845
 Arg Glu Tyr Phe
 850

<210> 121
 <211> 1905
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 121
 atgaagcata tttttattgt attaatgttt tccctgctgt ttagcttcgg gggatatgct 60
 caacaaacca ttagcagagc tccgcagggg ttgaccagc aacgtgccgg cattgcatcc 120
 ggtaaaagttg aaatcgtaac ctataaatcg aaaaccgtag gagtgaatcg ctctgcacgt 180
 gtttatacac cagccggatt ctcaaaaaag aagaaatatt ctgtgcttta tttattacat 240
 ggcattggag gcgacgaaga tgagtggtag aaaaacggcg ttcctcatat tattttcgac 300
 aacctgattg ccgacggcaa aatggaaccg atgattgtgg tactgcccac tggctcgccg 360
 atgaaaaacg accgtgccga aggaaatatt ttcgacaaag agaaaagtga agcctttgca 420
 acattcgaaa aagacctttt aaacgattta ataccgttta tcgaaaaaaa ataccctgta 480
 ttaaaaaccc gtgagtttcg cgccattgca ggattatcaa tgggcggcgg acaatcgctc 540
 aattttggac tgggaaatct cgacaaattt gcatgggtag gcggcttttc atcggccccc 600
 aataccaaaa tgcccgtga gttggttcca aacactcaaa aggcaacaga aatgcttaag 660
 ttgctttatg tgtcttggtg cgataaagac aatttaatgc aggttagtca gcgcaccac 720
 gattatctga aagccaataa agtacctcat attttcaggg ttattcctga tggttaccac 780
 gattttaatg tttggaaaga cgatttgtat cattacgtac aaatgctgtt taagcctgtg 840
 gtaatgcccg tagcagcagc tactttaaaa gatgcttata aagggaaatt cttcattgga 900
 actgccctta ataccctca aattttgggt accgctgttg atgaagtga tattgttaaa 960
 accattttca actccattgt tgccgaaaac tgtatgaaga gtggcccgat gcaaccacaa 1020
 gaagggaaat ttgagtttga cctggccgat aagttttagt agtttggagt taaaaacaat 1080
 atgcagatta ttggtcatat gcttatctgg cattcgcagg caccgccgtg gttttttacc 1140
 gacagcgaag gcaaggacgt atcgcccagag gtgcttaccg agcgcagtga aaaccatatt 1200
 tatactgttg ttggccgtta caaaggcaag gtgcacggat gggatgtggt gaatgaagcc 1260
 atagttagc atggagccta ccgaaacagt aaattctacc aaatactggg cgaagatttt 1320
 atcaaaactg cattccagtt tgctcatgaa gccgaccctg atgcagaatt gtactacaac 1380
 gattattccg aattttgttc tgccaaaaga gaaggcattg cccgcattgg gaagaaactc 1440
 aaagaccagg gcattagaat cgacggcggt ggatttcagt gccatattgg cctcgattat 1500
 ccaggcctgg atgaatacga aaaaaccatt caattaattg ccaacgaggg ggtaaaagta 1560
 atgataaccg aaatggaaat atcggtatta cccatgcccg actggcgcggt tgggtgctgag 1620
 atttcggcca gtttcgaata tcaacagaaa ttaaatccct acaccgaagg attgcccgat 1680

tcagtgaatg	ctcaattaga	acagcggtat	gtcgcactttt	tcacgctctt	ccttaaatat	1740
cacgaagtga	ttccaagagt	tacggtttgg	gggggttaacg	atggcaactc	atggaaaaac	1800
ggattcccg	tgcggtggaag	aaccgactac	ccattgttat	tcgaccggaa	aaatcagcct	1860
aaatcagctg	ttgccaaatt	aattgaactg	gctaatacaa	agtag		1905

<210> 122
 <211> 634
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(20)

<400> 122
 Met Lys His Ile Phe Ile Val Leu Ile Val Ser Leu Leu Phe Ser Phe
 1 5 10 15
 Gly Gly Tyr Ala Gln Gln Thr Ile Ser Arg Ala Pro Gln Gly Phe Asp
 20 25 30
 Gln Gln Arg Ala Gly Ile Ala Ser Gly Lys Val Glu Ile Val Thr Tyr
 35 40 45
 Lys Ser Lys Thr Val Gly Val Asn Arg Ser Ala Arg Val Tyr Thr Pro
 50 55 60
 Ala Gly Phe Ser Lys Lys Lys Tyr Pro Val Leu Tyr Leu Leu His
 65 70 75 80
 Gly Ile Gly Gly Asp Glu Asp Glu Trp Tyr Lys Asn Gly Val Pro His
 85 90 95
 Ile Ile Phe Asp Asn Leu Ile Ala Asp Gly Lys Met Glu Pro Met Ile
 100 105 110
 Val Val Leu Pro Asn Gly Arg Ala Met Lys Asn Asp Arg Ala Glu Gly
 115 120 125
 Asn Ile Phe Asp Lys Glu Lys Val Glu Ala Phe Ala Thr Phe Glu Lys
 130 135 140
 Asp Leu Leu Asn Asp Leu Ile Pro Phe Ile Glu Lys Lys Tyr Pro Val
 145 150 155 160
 Leu Lys Thr Arg Glu Phe Arg Ala Ile Ala Gly Leu Ser Met Gly Gly
 165 170 175
 Gly Gln Ser Leu Asn Phe Gly Leu Gly Asn Leu Asp Lys Phe Ala Trp
 180 185 190
 Val Gly Gly Phe Ser Ser Ala Pro Asn Thr Lys Met Pro Ala Glu Leu
 195 200 205
 Val Pro Asn Thr Gln Lys Ala Thr Glu Met Leu Lys Leu Leu Tyr Val
 210 215 220
 Ser Cys Gly Asp Lys Asp Asn Leu Met Gln Val Ser Gln Arg Thr His
 225 230 235 240
 Asp Tyr Leu Lys Ala Asn Lys Val Pro His Ile Phe Arg Val Ile Pro
 245 250 255
 Asp Gly Tyr His Asp Phe Asn Val Trp Lys Asp Asp Leu Tyr His Tyr
 260 265 270
 Val Gln Met Leu Phe Lys Pro Val Val Met Pro Val Ala Ala Thr
 275 280 285
 Leu Lys Asp Ala Tyr Lys Gly Lys Phe Phe Ile Gly Thr Ala Leu Asn
 290 295 300
 Thr Pro Gln Ile Leu Gly Thr Ala Val Asp Glu Val Asn Ile Val Lys
 305 310 315 320
 Thr His Phe Asn Ser Ile Val Ala Glu Asn Cys Met Lys Ser Gly Pro
 325 330 335
 Met Gln Pro Gln Glu Gly Lys Phe Glu Phe Asp Leu Ala Asp Lys Phe
 340 345 350
 Val Glu Phe Gly Val Lys Asn Asn Met Gln Ile Ile Gly His Thr Leu
 355 360 365
 Ile Trp His Ser Gln Ala Pro Arg Trp Phe Phe Thr Asp Ser Glu Gly
 370 375 380
 Lys Asp Val Ser Pro Glu Val Leu Thr Glu Arg Met Lys Asn His Ile
 385 390 395 400
 Tyr Thr Val Val Gly Arg Tyr Lys Gly Lys Val His Gly Trp Asp Val
 405 410 415

Val Asn Glu Ala Ile Val Asp Asp Gly Ser Tyr Arg Asn Ser Lys Phe
 420 425 430
 Tyr Gln Ile Leu Gly Glu Asp Phe Ile Lys Leu Ala Phe Gln Phe Ala
 435 440 445
 His Glu Ala Asp Pro Asp Ala Glu Leu Tyr Tyr Asn Asp Tyr Ser Glu
 450 455 460
 Phe Val Pro Ala Lys Arg Glu Gly Ile Ala Arg Met Val Lys Lys Leu
 465 470 475 480
 Lys Asp Gln Gly Ile Arg Ile Asp Gly Val Gly Phe Gln Cys His Ile
 485 490 495
 Gly Leu Asp Tyr Pro Gly Leu Asp Glu Tyr Glu Lys Thr Ile Gln Leu
 500 505 510
 Ile Ala Asn Glu Gly Val Lys Val Met Ile Thr Glu Met Glu Ile Ser
 515 520 525
 Val Leu Pro Met Pro Asp Trp Arg Val Gly Ala Glu Ile Ser Ala Ser
 530 535 540
 Phe Glu Tyr Gln Gln Lys Leu Asn Pro Tyr Thr Glu Gly Leu Pro Asp
 545 550 555 560
 Ser Val Asn Ala Gln Leu Glu Gln Arg Tyr Val Asp Phe Phe Thr Leu
 565 570 575
 Phe Leu Lys Tyr His Glu Val Ile Pro Arg Val Thr Val Trp Gly Val
 580 585 590
 Asn Asp Gly Asn Ser Trp Lys Asn Gly Phe Pro Val Arg Gly Arg Thr
 595 600 605
 Asp Tyr Pro Leu Leu Phe Asp Arg Lys Asn Gln Pro Lys Ser Ala Val
 610 615 620
 Ala Lys Leu Ile Glu Leu Ala Asn Thr Lys
 625 630

<210> 123
 <211> 1200
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 123
 atgatcgttg gattctcgtt tatgctgctg cttccttttag ggatgacgaa tgcattggca 60
 aaaacggaac cagcgtacgc taaaaagccg cgaatcagcg cattgcacgc ccctcaattg 120
 gatcagcgct acaaaagattc cttcactatt gggcgcgccg ttgaacctta tcagttgcaa 180
 aacgaaaaag acgtccaaat gctgaaacgc catitttaaca gcattgtcgc tgagaacggt 240
 atgaaaccga tcaacatcca acccgaagaa ggaaagtcca attttgctga ggcggatcaa 300
 atcgtccgat ttgctaaaaa acatcatatg gatattcgtt tccatacact cgtttggcac 360
 agccaagtac ctcaatgggt ctttcttgac aagggaaggca agccgatggt caatgaaacg 420
 gatccggcaa agcgcgaaca aaataaacag ctgttactga aacggctcga aatccatatt 480
 aaaacgattg tcgaacggtg taaagacgac atcaaatatt gggacgtcgt gaacgaggta 540
 gtcgggggatg atggaaaatt gcgcaattcc ccgtgggtatc aaatcgccgg catcgattat 600
 atcaaggtag cattccaaac ggcgagaaca tatggcggca acaagattaa actgtacatc 660
 aacgattaca ataccgaagt ggaaccgaag cgaagcgctc ttataactt agtgaaacaa 720
 ttaaaagaag aaggcgttcc cattgacggg attggccacc agtcccacat ccaaattggc 780
 tggccttctg aagaagaaat cgaaaaaacg atcaacatgt ttgccgatct aggggttagac 840
 aatcaaatta cggagctgga tgtgagcatg tacggctggc cgccgcgcgc ctaccgctcg 900
 tatgacgcca ttccggaaca aaagtttttg gaccaagcgg ctcgctatga ccgattgttt 960
 aagctgtacg aaaaacttgg cgataaaatc agcaacgtca ccttctgggg catcgccgac 1020
 aaccatacgt ggctcgacag ccgtgcggat gtgtactatg acgccaacgg gaatgttgtg 1080
 gttgacccga acgctccgta cgcaaaagtg gaaaaagggg aaggaaaaga tgcgccgttt 1140
 ctgttcgacc ccgaatacca cgtaaaacct gcgtattggg ccattatcga tcataagtga 1200

<210> 124
 <211> 399
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(20)

<400> 124
 Met Ile Val Gly Phe Ser Phe Met Leu Leu Leu Pro Leu Gly Met Thr
 1 5 10 15
 Asn Ala Leu Ala Lys Thr Glu Pro Ala Tyr Ala Lys Lys Pro Arg Ile
 20 25 30
 Ser Ala Leu His Ala Pro Gln Leu Asp Gln Arg Tyr Lys Asp Ser Phe
 35 40 45
 Thr Ile Gly Ala Ala Val Glu Pro Tyr Gln Leu Gln Asn Glu Lys Asp
 50 55 60
 Val Gln Met Leu Lys Arg His Phe Asn Ser Ile Val Ala Glu Asn Val
 65 70 75 80
 Met Lys Pro Ile Asn Ile Gln Pro Glu Glu Gly Lys Phe Asn Phe Ala
 85 90 95
 Glu Ala Asp Gln Ile Val Arg Phe Ala Lys Lys His His Met Asp Ile
 100 105 110
 Arg Phe His Thr Leu Val Trp His Ser Gln Val Pro Gln Trp Phe Phe
 115 120 125
 Leu Asp Lys Glu Gly Lys Pro Met Val Asn Glu Thr Asp Pro Ala Lys
 130 135 140
 Arg Glu Gln Asn Lys Gln Leu Leu Lys Arg Leu Glu Ile His Ile
 145 150 155 160
 Lys Thr Ile Val Glu Arg Tyr Lys Asp Asp Ile Lys Tyr Trp Asp Val
 165 170 175
 Val Asn Glu Val Val Gly Asp Asp Gly Lys Leu Arg Asn Ser Pro Trp
 180 185 190
 Tyr Gln Ile Ala Gly Ile Asp Tyr Ile Lys Val Ala Phe Gln Thr Ala
 195 200 205
 Arg Thr Tyr Gly Gly Asn Lys Ile Lys Leu Tyr Ile Asn Asp Tyr Asn
 210 215 220
 Thr Glu Val Glu Pro Lys Arg Ser Ala Leu Tyr Asn Leu Val Lys Gln
 225 230 235 240
 Leu Lys Glu Glu Gly Val Pro Ile Asp Gly Ile Gly His Gln Ser His
 245 250 255
 Ile Gln Ile Gly Trp Pro Ser Glu Glu Ile Glu Lys Thr Ile Asn
 260 265 270
 Met Phe Ala Asp Leu Gly Leu Asp Asn Gln Ile Thr Glu Leu Asp Val
 275 280 285
 Ser Met Tyr Gly Trp Pro Pro Arg Ala Tyr Pro Ser Tyr Asp Ala Ile
 290 295 300
 Pro Glu Gln Lys Phe Leu Asp Gln Ala Ala Arg Tyr Asp Arg Leu Phe
 305 310 315 320
 Lys Leu Tyr Glu Lys Leu Gly Asp Lys Ile Ser Asn Val Thr Phe Trp
 325 330 335
 Gly Ile Ala Asp Asn His Thr Trp Leu Asp Ser Arg Ala Asp Val Tyr
 340 345 350
 Tyr Asp Ala Asn Gly Asn Val Val Val Asp Pro Asn Ala Pro Tyr Ala
 355 360 365
 Lys Val Glu Lys Gly Lys Gly Lys Asp Ala Pro Phe Leu Phe Asp Pro
 370 375 380
 Glu Tyr His Val Lys Pro Ala Tyr Trp Ala Ile Ile Asp His Lys
 385 390 395

<210> 125

<211> 1089

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 125

atgttgacga	ccccgacaac	tcaagatcat	gtccccgtgc	tcaaggacgc	tttcaaaggc	60
aagctcctca	ttggagccgt	gctcgggttac	gatgctctcc	aggggaagga	cccgtgagt	120
gagaaaattg	cgaccactca	cttcgatgct	ctcactgctg	aaaacagcat	gaagccggct	180
ctcgtgaac	ccaaagaggg	cgagtttgat	ttcgtgatg	gagatcgtct	ccttgaaatc	240
gcgcagcaat	gcggcgctac	tgcaatcggc	catactctgc	tctggcacca	acaaacgcca	300
cgctggtttt	ttgaagggcc	agatggtcag	cctgctgacc	gtgagttggc	cctggcacgc	360
atgaggaagc	acatttccac	tctcgttggg	cgctataaag	gtcgcattaa	acaatgggat	420

gtggtgaatg	aggcgattag	cgatgcagag	ggcagagtact	taagaccaa	gagcccctgg	480
ttcaaagccg	ttggagagga	tcacatcgcg	catgctttcc	aggcagcaca	tgaagctgat	540
cccgatgcc	tccttatcta	taacgactac	aacatcgagc	aggagtacaa	gcgcccgaag	600
gcgatacgcc	tactgaggtc	attacttgag	caggacgttc	ccattcatgc	cgtagggcatt	660
cagggccatt	ggcgtatgga	cactctgaat	gttgccgaaa	tcgaagaagc	tatcgaagaa	720
tttgctgcgc	tgggtctcaa	ggtcatgatc	accgagcttg	atatcagcgt	gctaccgaca	780
aagtatcagg	gagccgatct	cgctactcgg	gaagaattga	cgctgaaat	caatccctat	840
acggaggaac	tacctgagga	cgttgcccg	caacatgccg	agtgttatcg	gcagggtcttc	900
gaaatgttcc	tcgcccacaa	ggatgccatt	agccgtgtca	cgctctgggg	cattcacgat	960
ggcagatcat	ggttcaacaa	ctttccggtc	agggggcgca	cagactatcc	tctgctattc	1020
gaccgggaat	gtaaccccaa	gccagcgttt	ttcgccgtct	tgaaagctgc	gcaagaccag	1080
ccacaatga						1089

<210> 126
 <211> 362
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 126
 Met Leu Thr Thr Pro Thr Thr Gln Asp His Val Pro Val Leu Lys Asp
 1 5 10 15
 Ala Phe Lys Gly Lys Leu Leu Ile Gly Ala Val Leu Gly Tyr Asp Ala
 20 25 30
 Leu Gln Gly Lys Asp Pro Leu Ser Glu Lys Ile Ala Thr Thr His Phe
 35 40 45
 Asp Ala Leu Thr Ala Glu Asn Ser Met Lys Pro Ala Leu Val Gln Pro
 50 55 60
 Lys Glu Gly Glu Phe Asp Phe Ala Asp Gly Asp Arg Leu Leu Glu Ile
 65 70 75 80
 Ala Gln Gln Cys Gly Ala Thr Ala Ile Gly His Thr Leu Leu Trp His
 85 90 95
 Gln Gln Thr Pro Arg Trp Phe Phe Glu Gly Pro Asp Gly Gln Pro Ala
 100 105 110
 Asp Arg Glu Leu Ala Leu Ala Arg Met Arg Lys His Ile Ser Thr Leu
 115 120 125
 Val Gly Arg Tyr Lys Gly Arg Ile Lys Gln Trp Asp Val Val Asn Glu
 130 135 140
 Ala Ile Ser Asp Ala Glu Gly Glu Tyr Leu Arg Pro Lys Ser Pro Trp
 145 150 155 160
 Phe Lys Ala Val Gly Glu Asp His Ile Ala His Ala Phe Gln Ala Ala
 165 170 175
 His Glu Ala Asp Pro Asp Ala Ile Leu Ile Tyr Asn Asp Tyr Asn Ile
 180 185 190
 Glu Gln Glu Tyr Lys Arg Pro Lys Ala Ile Arg Leu Leu Arg Ser Leu
 195 200 205
 Leu Glu Gln Asp Val Pro Ile His Ala Val Gly Ile Gln Gly His Trp
 210 215 220
 Arg Met Asp Thr Leu Asn Val Ala Glu Ile Glu Glu Ala Ile Glu Glu
 225 230 235 240
 Phe Ala Ala Leu Gly Leu Lys Val Met Ile Thr Glu Leu Asp Ile Ser
 245 250 255
 Val Leu Pro Thr Lys Tyr Gln Gly Ala Asp Leu Ala Thr Arg Glu Glu
 260 265 270
 Leu Thr Pro Glu Ile Asn Pro Tyr Thr Glu Glu Leu Pro Glu Asp Val
 275 280 285
 Ala Arg Gln His Ala Glu Cys Tyr Arg Gln Val Phe Glu Met Phe Leu
 290 295 300
 Arg His Lys Asp Ala Ile Ser Arg Val Thr Leu Trp Gly Ile His Asp
 305 310 315 320
 Gly Arg Ser Trp Phe Asn Asn Phe Pro Val Arg Gly Arg Thr Asp Tyr
 325 330 335
 Pro Leu Leu Phe Asp Arg Glu Cys Asn Pro Lys Pro Ala Phe Phe Ala
 340 345 350
 Val Leu Lys Ala Ala Gln Asp Gln Pro Gln
 355 360

<210> 127
 <211> 960
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 127
 gtggatctcg ctgagaaatg cggcatatat attggtgcag cggttgaacc cggatattta 60
 attatcaggg aatacgctga gattttatcc cgcgaattta acgtggtaac cgcggaaaat 120
 gcattaaaaa ttgaagctat tcatccgcag cgtggagtat attcatttga aggtgcagat 180
 gcaatagttc gatttgcaga aactcatgga atgaaggttc gtggacatac acttgtttgg 240
 caccagcagc ttcctgcatg gataacttct ggaagttagc ctggggagga gtggaagaat 300
 attctccgtg agcatgtaat gagcgttggt ggacgatata agggccaaat atatgcatgg 360
 gatgtggtta acgaagcaat attagataac ggttcattaa gagataatgt ttggtttaga 420
 aatgtaggtc cagaatatat tgagtcagcc tttagatggg ctcatgaagc tgaccctaac 480
 gctcttctct tctataatga ttatgaagct gaggacttga atgataagtc gcatgctgtt 540
 tataacctgg ttaagagttt acttgagaaa ggtgtaccga tacatggcgt aggtattacag 600
 atgcatatta acgtagaaaa tccgccgaaa ccggaagatg ttgcagcaaa cattaacagt 660
 ctaaatgatc tgggcttgat tgtccacata acggaatgg atgtgcgcat tagaacccca 720
 ccatcaaatg aagaatctcat taaacaagca gaaatttacc gtgatataat aagagtttgt 780
 ctttcatcag aaaaatgcac agcattcatt atgtggggat ttactgaccg ctattcatgg 840
 ataccaaatt acttcagcgg ctacggttca gctttaatat tcgatgagca atataagccc 900
 aaactagcat attactatat acttcggaca ttcatcgaaa aactaggcat taaaggtaa 960

<210> 128
 <211> 319
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 128
 Val Asp Leu Ala Glu Lys Cys Gly Ile Tyr Ile Gly Ala Ala Val Glu
 1 5 10 15
 Pro Gly Tyr Leu Ile Ile Arg Glu Tyr Ala Glu Ile Leu Ser Arg Glu
 20 25 30
 Phe Asn Val Val Thr Ala Glu Asn Ala Leu Lys Phe Glu Ala Ile His
 35 40 45
 Pro Gln Arg Gly Val Tyr Ser Phe Glu Gly Ala Asp Ala Ile Val Arg
 50 55 60
 Phe Ala Glu Thr His Gly Met Lys Val Arg Gly His Thr Leu Val Trp
 65 70 75 80
 His Gln Gln Leu Pro Ala Trp Ile Thr Ser Gly Ser Tyr Ala Trp Glu
 85 90 95
 Glu Trp Lys Asn Ile Leu Arg Glu His Val Met Ser Val Val Gly Arg
 100 105 110
 Tyr Lys Gly Gln Ile Tyr Ala Trp Asp Val Val Asn Glu Ala Ile Leu
 115 120 125
 Asp Asn Gly Ser Leu Arg Asp Asn Val Trp Phe Arg Asn Val Gly Pro
 130 135 140
 Glu Tyr Ile Glu Ser Ala Phe Arg Trp Ala His Glu Ala Asp Pro Asn
 145 150 155 160
 Ala Leu Leu Phe Tyr Asn Asp Tyr Glu Ala Glu Asp Leu Asn Asp Lys
 165 170 175
 Ser His Ala Val Tyr Asn Leu Val Lys Ser Leu Leu Glu Lys Gly Val
 180 185 190
 Pro Ile His Gly Val Gly Leu Gln Met His Ile Asn Val Glu Asn Pro
 195 200 205
 Pro Lys Pro Glu Asp Val Ala Asn Ile Lys Arg Leu Asn Asp Leu
 210 215 220
 Gly Leu Ile Val His Ile Thr Glu Met Asp Val Arg Ile Arg Thr Pro
 225 230 235 240
 Pro Ser Asn Glu Asp Leu Ile Lys Gln Ala Glu Ile Tyr Arg Asp Ile
 245 250 255
 Leu Arg Val Cys Leu Ser Ser Glu Lys Cys Thr Ala Phe Ile Met Trp
 260 265 270

Gly Phe Thr Asp Arg Tyr Ser Trp Ile Pro Asn Tyr Phe Ser Gly Tyr
 275 280 285
 Gly Ser Ala Leu Ile Phe Asp Glu Gln Tyr Lys Pro Lys Leu Ala Tyr
 290 295 300
 Tyr Tyr Ile Leu Arg Thr Phe Ile Glu Lys Leu Gly Ile Lys Gly
 305 310 315

<210> 129
 <211> 3021
 <212> DNA
 <213> Bacteria

<400> 129
 atggtaataa atcgctccag tgcgagtgc ggtgcgtatt cggaaaaagg tttctatctc 60
 gacggtggtg tagaatacaa gtacagtgtt tttgtaaaac acaacgggac cggcaccgaa 120
 actttcaaac ttctgtgtc ctatttggat tcggaacag aagaagaaa taaggaaagta 180
 attgcaacaa aggatgttgt ggccggagaa ttccggcaaa atacaaagca 240
 cccaaaactg cagtgaatat tactttgtca attacaaccg acagcactgt agatttcatt 300
 tttgacgatg taaccataac ccgtaaagga atggctgagg caaacacagt atatgcagca 360
 aacgctgtgc tgaagatat gtatgcaaac tatttcagag ttggttcggt acttaactcc 420
 ggaacggtaa acaattcatc aataaaggcc ttgattttaa gagagttaa cagtattacc 480
 tgtgaaaatg aaatgaagcc tgatgccaca ctggttcaat caggatcaac caatacaaat 540
 atcagggttt ctcttaatcg tgcagcaagt attttaaact tctgtgcaca aaataatata 600
 gccgtcagag gtcatacact ggtttggcac agccagacac ctcaatggtt tttcaaagac 660
 aatttccagg acaacggaaa ctgggtttcc caatcagtta tggaccagcg tttgaaagc 720
 tacataaaaa atatgtttgc tgaaatccaa agacagtatc cgtctttgaa tctttatgcc 780
 tatgacgttg taaatgaggg agtaagtgat gatgcaaaca ggaccagata ttatggcggg 840
 gcgagggaac ctggatacgg aaatggtaga tctccatggg ttcagatcta cggagacaac 900
 aaattttatt agaaagcatt tacatatgca agaaaatatg ctccggcaaa ttgtaagctt 960
 tactacaacg attacaacga atattgggat actgtattgc ctcaatttgt 1020
 gcaaacttgt acaacaaggg ctgtctgac ggtgtgggaa tgcagtccca tattaatgcg 1080
 gatatgaatg gattctcagg tatacaaaat tataaagcag ctttgcagaa atatataaat 1140
 atcggttgtg atgtccaaat taccgagctt gatattagta cagaaaacgg caaatttagc 1200
 ttacagcagc aggtgataa atataaagct gttttccagg cagctgttga tataaacaga 1260
 acctccagca aaggaaaagg taccgctgtc tgtgtatggg gacctaataga cgccaatact 1320
 tggctcgggt cacaanaatgc acctcttttg tttaacgcaa acaatcaacc gaaaccggca 1380
 tacaatgcgg ttgcatccat tattcctcag tccgaatggg gcgacggtaa caatccggcc 1440
 ggcggcggag gaggaggcaa accggaagag ccggaatgcaa acggatatta ttatcatgac 1500
 acttttgaag gtaacgttagg acagtggaca gccagaggac ctgcggaagt tctgcttagc 1560
 ggaagaacgg cttacaaggg ttcagaatca ctcttggtta ggaaccgtac ggcagcatgg 1620
 aacggagcac aacgggcgct gaatcccaga acgtttgttc ccggaaacac atattgtttc 1680
 agcgtagtgg catcgtttat tgaagggtcg tcttccacaa cattctgcat gaagctgcaa 1740
 tacgtagacg gaagcggcac tcaacggtat gataccatag atatgaaaac tgtgggtcca 1800
 aatcagttggg ttcacctgta caatccgcaa ttccgatgc aacagatatg 1860
 tatgtttatg tggaaacagc ggatgacacc attaatctct acatagatga ggcaatcggg 1920
 gcggttgccg gaactgtaat cgaaggacct gctccacagc ctacacagcc tccggtactg 1980
 cttggcgatg taaacgggtga tggaaaccatt aactcaactg acttgacaat gttaaagaga 2040
 agcgtgttga gggcaatcac ccttaccgag gatgcaaagg cttagagcaga cgttgacaag 2100
 aatggatcga taaacagcac tgatgtttta cttctttcac gctacccttt aagagtaatc 2160
 gacaaatttc ctgtagcaga aaatccttct tcttctttta aatatgagtc ggccgtgcaa 2220
 tatcggccgg ctctgatttc ttattttaaac ccttgtccgc aggcgggaag aattgtcaag 2280
 gaaacataa caggaataaa cggaactaag agtcttaatg tatactttcc atacggttat 2340
 gatccgaaca aaaaatataa cattttctac cttatgcatg gcggcgggtg aaatgagaat 2400
 acgattttca gcaacgatgt taaattgcaa aatatccttg accacgcgat tatgaacggt 2460
 gaacttgagc ctttgattgt agtaacaccc actttcaacg gcggaactg cacggcccaa 2520
 aacttttatc aggaattcag gcaaaatgtc attccttttg tggaaagcaa gtactctact 2580
 tatgcagaat caacaacccc acaggaata gccgcttcaa gaatgcacag aggtttcggc 2640
 ggattctcaa tgggaggatt gacaacatgg tatgtaatgg ttaactgcct tgattacgtt 2700
 gcatatttta tgcctttaag cggtgactac tggatggaa acagtccgca ggataaggct 2760
 aattcaattg ctgaagcaat taacagatcc ggactttcaa agagggagta tttcgtattt 2820
 gcggccaccg gttccgagga tattgcatat gctaataatga atcctcaaat tgaagctatg 2880
 aaggctttgc cgcattttga ttatacttcg aaggttaattt ttactttctt 2940
 gtatgctccg gcgccactca ctggtgggga tacgtaagac attatattta tgatgcactt 3000
 ccatatttct tccatgaatg a 3021

<210> 130
 <211> 1006
 <212> PRT
 <213> Bacteria

<400> 130
 Met Val Ile Asn Arg Ser Ser Ala Ser Asp Gly Ala Tyr Ser Glu Lys
 1 5 10 15
 Gly Phe Tyr Leu Asp Gly Gly Val Glu Tyr Lys Tyr Ser Val Phe Val
 20 25 30
 Lys His Asn Gly Thr Gly Thr Glu Thr Phe Lys Leu Ser Val Ser Tyr
 35 40 45
 Leu Asp Ser Glu Thr Glu Glu Glu Asn Lys Glu Val Ile Ala Thr Lys
 50 55 60
 Asp Val Val Ala Gly Glu Trp Thr Glu Ile Ser Ala Lys Tyr Lys Ala
 65 70 75 80
 Pro Lys Thr Ala Val Asn Ile Thr Leu Ser Ile Thr Thr Asp Ser Thr
 85 90 95
 Val Asp Phe Ile Phe Asp Asp Val Thr Ile Thr Arg Lys Gly Met Ala
 100 105 110
 Glu Ala Asn Thr Val Tyr Ala Ala Asn Ala Val Leu Lys Asp Met Tyr
 115 120 125
 Ala Asn Tyr Phe Arg Val Gly Ser Val Leu Asn Ser Gly Thr Val Asn
 130 135 140
 Asn Ser Ser Ile Lys Ala Leu Ile Leu Arg Glu Phe Asn Ser Ile Thr
 145 150 155 160
 Cys Glu Asn Glu Met Lys Pro Asp Ala Thr Leu Val Gln Ser Gly Ser
 165 170 175
 Thr Asn Thr Asn Ile Arg Val Ser Leu Asn Arg Ala Ala Ser Ile Leu
 180 185 190
 Asn Phe Cys Ala Gln Asn Asn Ile Ala Val Arg Gly His Thr Leu Val
 195 200 205
 Trp His Ser Gln Thr Pro Gln Trp Phe Phe Lys Asp Asn Phe Gln Asp
 210 215 220
 Asn Gly Asn Trp Val Ser Gln Ser Val Met Asp Gln Arg Leu Glu Ser
 225 230 235 240
 Tyr Ile Lys Asn Met Phe Ala Glu Ile Gln Arg Gln Tyr Pro Ser Leu
 245 250 255
 Asn Leu Tyr Ala Tyr Asp Val Val Asn Glu Ala Val Ser Asp Asp Ala
 260 265 270
 Asn Arg Thr Arg Tyr Tyr Gly Gly Ala Arg Glu Pro Gly Tyr Gly Asn
 275 280 285
 Gly Arg Ser Pro Trp Val Gln Ile Tyr Gly Asp Asn Lys Phe Ile Glu
 290 295 300
 Lys Ala Phe Thr Tyr Ala Arg Lys Tyr Ala Pro Ala Asn Cys Lys Leu
 305 310 315 320
 Tyr Tyr Asn Asp Tyr Asn Glu Tyr Trp Asp His Lys Arg Asp Cys Ile
 325 330 335
 Ala Ser Ile Cys Ala Asn Leu Tyr Asn Lys Gly Leu Leu Asp Gly Val
 340 345 350
 Gly Met Gln Ser His Ile Asn Ala Asp Met Asn Gly Phe Ser Gly Ile
 355 360 365
 Gln Asn Tyr Lys Ala Ala Leu Gln Lys Tyr Ile Asn Ile Gly Cys Asp
 370 375 380
 Val Gln Ile Thr Glu Leu Asp Ile Ser Thr Glu Asn Gly Lys Phe Ser
 385 390 395 400
 Leu Gln Gln Gln Ala Asp Lys Tyr Lys Ala Val Phe Gln Ala Ala Val
 405 410 415
 Asp Ile Asn Arg Thr Ser Ser Lys Gly Lys Val Thr Ala Val Cys Val
 420 425 430
 Trp Gly Pro Asn Asp Ala Asn Thr Trp Leu Gly Ser Gln Asn Ala Pro
 435 440 445
 Leu Leu Phe Asn Ala Asn Asn Gln Pro Lys Pro Ala Tyr Asn Ala Val
 450 455 460
 Ala Ser Ile Ile Pro Gln Ser Glu Trp Gly Asp Gly Asn Asn Pro Ala
 465 470 475 480
 Gly Gly Gly Gly Gly Gly Lys Pro Glu Glu Pro Asp Ala Asn Gly Tyr
 485 490 495
 Tyr Tyr His Asp Thr Phe Glu Gly Ser Val Gly Gln Trp Thr Ala Arg
 500 505 510
 Gly Pro Ala Glu Val Leu Leu Ser Gly Arg Thr Ala Tyr Lys Gly Ser
 515 520 525
 Glu Ser Leu Leu Val Arg Asn Arg Thr Ala Ala Trp Asn Gly Ala Gln

	530					535					540				
Arg 545	Ala	Leu	Asn	Pro	Arg 550	Thr	Phe	Val	Pro	Gly 555	Asn	Thr	Tyr	Cys	Phe 560
Ser	Val	Val	Ala	Ser 565	Phe	Ile	Glu	Gly	Ala 570	Ser	Ser	Thr	Thr	Phe 575	Cys
Met	Lys	Leu	Gln 580	Tyr	Val	Asp	Gly	Ser 585	Gly	Thr	Gln	Arg	Tyr 590	Asp	Thr
Ile	Asp	Met 595	Lys	Thr	Val	Gly	Pro 600	Asn	Gln	Trp	Val	His 605	Leu	Tyr	Asn
Pro	Gln 610	Tyr	Arg	Ile	Pro	Ser 615	Asp	Ala	Thr	Asp	Met 620	Tyr	Val	Tyr	Val
Glu 625	Thr	Ala	Asp	Asp	Thr 630	Ile	Asn	Phe	Tyr	Ile 635	Asp	Glu	Ala	Ile	Gly 640
Ala	Val	Ala	Gly	Thr 645	Val	Ile	Glu	Gly	Pro 650	Ala	Pro	Gln	Pro	Thr 655	Gln
Pro	Pro	Val	Leu 660	Leu	Gly	Asp	Val	Asn 665	Gly	Asp	Gly	Thr	Ile 670	Asn	Ser
Thr	Asp	Leu 675	Thr	Met	Leu	Lys	Arg 680	Ser	Val	Leu	Arg	Ala 685	Ile	Thr	Leu
Thr	Asp 690	Asp	Ala	Lys	Ala	Arg 695	Ala	Asp	Val	Asp	Lys 700	Asn	Gly	Ser	Ile
Asn 705	Ser	Thr	Asp	Val	Leu 710	Leu	Leu	Ser	Arg	Tyr 715	Leu	Leu	Arg	Val	Ile 720
Asp	Lys	Phe	Pro	Val 725	Ala	Glu	Asn	Pro	Ser 730	Ser	Ser	Phe	Lys	Tyr 735	Glu
Ser	Ala	Val	Gln 740	Tyr	Arg	Pro	Ala	Pro 745	Asp	Ser	Tyr	Leu	Asn 750	Pro	Cys
Pro	Gln	Ala 755	Gly	Arg	Ile	Val	Lys 760	Glu	Thr	Tyr	Thr	Gly 765	Ile	Asn	Gly
Thr	Lys 770	Ser	Leu	Asn	Val	Tyr 775	Leu	Pro	Tyr	Gly	Tyr 780	Asp	Pro	Asn	Lys
Lys 785	Tyr	Asn	Ile	Phe	Tyr 790	Leu	Met	His	Gly	Gly 795	Gly	Glu	Asn	Glu	Asn 800
Thr	Ile	Phe	Ser	Asn 805	Asp	Val	Lys	Leu	Gln 810	Asn	Ile	Leu	Asp	His 815	Ala
Ile	Met	Asn	Gly 820	Glu	Leu	Glu	Pro	Leu 825	Ile	Val	Val	Thr	Pro 830	Thr	Phe
Asn	Gly	Gly 835	Asn	Cys	Thr	Ala	Gln 840	Asn	Phe	Tyr	Gln	Glu 845	Phe	Arg	Gln
Asn	Val 850	Ile	Pro	Phe	Val	Glu 855	Ser	Lys	Tyr	Ser	Thr 860	Tyr	Ala	Glu	Ser
Thr 865	Thr	Pro	Gln	Gly	Ile 870	Ala	Ala	Ser	Arg	Met 875	His	Arg	Gly	Phe	Gly 880
Gly	Phe	Ser	Met	Gly 885	Gly	Leu	Thr	Thr	Trp 890	Tyr	Val	Met	Val	Asn 895	Cys
Leu	Asp	Tyr	Val 900	Ala	Tyr	Phe	Met	Pro 905	Leu	Ser	Gly	Asp	Tyr 910	Trp	Tyr
Gly	Asn	Ser 915	Pro	Gln	Asp	Lys	Ala 920	Asn	Ser	Ile	Ala	Glu 925	Ala	Ile	Asn
Arg	Ser 930	Gly	Leu	Ser	Lys	Arg 935	Glu	Tyr	Phe	Val	Phe 940	Ala	Ala	Thr	Gly
Ser 945	Glu	Asp	Ile	Ala	Tyr 950	Ala	Asn	Met	Asn	Pro 955	Gln	Ile	Glu	Ala	Met 960
Lys	Ala	Leu	Pro	His 965	Phe	Asp	Tyr	Thr	Ser 970	Asp	Phe	Ser	Lys	Gly 975	Asn
Phe	Tyr	Phe	Leu 980	Val	Ala	Pro	Gly	Ala 985	Thr	His	Trp	Trp	Gly 990	Tyr	Val
Arg	His	Tyr 995	Ile	Tyr	Asp	Ala	Leu 1000	Pro	Tyr	Phe	Phe	His 1005	Glu		

<210>	131
<211>	1218
<212>	DNA
<213>	Unknown

<220>
<223> Obtained from an environmental sample

<400> 131

atgccgatca	tccgaaccct	atcgagttac	atgcgaaatc	atcaagcgat	ctaccgtcag	60
ctcctcacgc	tggccgcgcg	cgtcacgctg	gcgggcgcg	ccaccgcgga	ggaagaagcc	120
accctgcgcg	gggtttacga	aaaggacttc	accatcggcg	tggccatgaa	cgggggccag	180
gcctccggcc	gcaatgccgc	cgccggcgag	atcatcgga	agcagttctc	ctcgtccacc	240
gcggagaacg	acatgaagtg	gcagatgata	cacccccagg	aggggtcaata	ccgcttcgaa	300
acgtccgacg	cctacgtcgc	gttcgcggaa	aagcacaaga	tggaaagtc	cggccacacc	360
ctcgtgtggc	acagccagac	cccgcagtgg	gtcttcagg	gtgaaaacgg	ccagcccgcc	420
accaagggaag	agctgctcaa	gcgcattgcg	gaccacatcc	acgccgtggc	cggccgttac	480
aagggcaaga	tcaagggctg	ggacgtcgtc	aacgaagcgc	tctccgacgg	cggggacgac	540
attctccgcc	agtccccctg	gcgccgcata	atcggcgacg	acttcacgca	ctacgccttc	600
cgctacgcc	aggaagccgc	cccggatgcc	gagctctact	acaacgacta	caacctcgag	660
atccccgcga	agcgcgcaa	ttgcatcacg	ctgggtcaagg	gcatgctcga	gcgcggcggtg	720
ccgatccgacg	gcctcggcac	ccagtgcgac	ttccagctcg	gctttccctc	cttgagcgac	780
gtggaagcca	ccatcaagga	attcgccgcc	ctgggcatga	aggtgatgat	caccgagctc	840
gacgtggatg	tcctgccccg	caacaacccc	ggggtcgccc	acatcgccaa	ccgcgaacag	900
ggagccaacc	cctacaccga	aggccttcgg	gacgacgtgc	aggaaaagct	cgcaagcgcg	960
tacgaggaca	tcttcggcat	ctacctgaag	taccgcgacc	acgtcacccg	cgtaaccttc	1020
tggggcctgg	atgacggcat	gacctggctg	aacggcttcc	cgggtccgcg	ccgcaccaac	1080
caccctctgc	tctacgaccg	gcagctcaat	gccaaagccg	ccttcacgcg	cctcgtcaag	1140
ctgggtcagg	aagagcgtcc	ggaagccgcc	aaggtcgagg	tccagaagat	cgaagcgaag	1200
aaagaagagg	cgaagtaa					1218

<210> 132
 <211> 405
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(26)

<400> 132

Met	Pro	Ile	Ile	Arg	Thr	Leu	Ser	Ser	Tyr	Met	Arg	Asn	His	Gln	Ala
1				5					10					15	
Ile	Tyr	Arg	Gln	Leu	Leu	Thr	Leu	Ala	Ala	Ala	Val	Thr	Leu	Ala	Gly
			20					25					30		
Ala	Ala	Thr	Ala	Glu	Glu	Glu	Ala	Thr	Leu	Arg	Gly	Val	Tyr	Glu	Lys
		35					40					45			
Asp	Phe	Thr	Ile	Gly	Val	Ala	Met	Asn	Gly	Gly	Gln	Ala	Ser	Gly	Arg
	50				55						60				
Asn	Ala	Ala	Ala	Gly	Glu	Ile	Ile	Gly	Lys	Gln	Phe	Ser	Ser	Leu	Thr
65					70					75				80	
Ala	Glu	Asn	Asp	Met	Lys	Trp	Gln	Met	Ile	His	Pro	Gln	Glu	Gly	Gln
			85						90					95	
Tyr	Arg	Phe	Glu	Thr	Ser	Asp	Ala	Tyr	Val	Ala	Phe	Ala	Glu	Lys	His
			100					105					110		
Lys	Met	Glu	Val	Ile	Gly	His	Thr	Leu	Val	Trp	His	Ser	Gln	Thr	Pro
		115					120					125			
Gln	Trp	Val	Phe	Gln	Gly	Glu	Asn	Gly	Gln	Pro	Ala	Thr	Lys	Glu	Glu
	130					135					140				
Leu	Leu	Lys	Arg	Met	Arg	Asp	His	Ile	His	Ala	Val	Ala	Gly	Arg	Tyr
145					150				155					160	
Lys	Gly	Lys	Ile	Lys	Gly	Trp	Asp	Val	Val	Asn	Glu	Ala	Leu	Ser	Asp
			165						170					175	
Gly	Gly	Asp	Asp	Ile	Leu	Arg	Gln	Ser	Pro	Trp	Arg	Arg	Ile	Ile	Gly
		180					185						190		
Asp	Asp	Phe	Ile	Asp	Tyr	Ala	Phe	Arg	Tyr	Ala	Lys	Glu	Ala	Ala	Pro
		195				200					205				
Asp	Ala	Glu	Leu	Tyr	Tyr	Asn	Asp	Tyr	Asn	Leu	Glu	Ile	Pro	Arg	Lys
	210					215					220				
Arg	Ala	Asn	Cys	Ile	Thr	Leu	Val	Lys	Gly	Met	Leu	Glu	Arg	Gly	Val
225					230					235				240	
Pro	Ile	Asp	Gly	Ile	Gly	Thr	Gln	Ser	His	Phe	Gln	Leu	Gly	Phe	Pro
		245							250					255	
Ser	Leu	Asp	Asp	Val	Glu	Ala	Thr	Ile	Lys	Glu	Phe	Ala	Ala	Leu	Gly
		260					265					270			
Met	Lys	Val	Met	Ile	Thr	Glu	Leu	Asp	Val	Asp	Val	Leu	Pro	Arg	Asn

275 280 285
 Asn Pro Gly Val Ala Asp Ile Ala Asn Arg Glu Gln Gly Ala Asn Pro
 290 295 300
 Tyr Thr Glu Gly Leu Pro Asp Asp Val Gln Glu Lys Leu Ala Lys Arg
 305 310 315 320
 Tyr Glu Asp Ile Phe Arg Ile Tyr Leu Lys Tyr Arg Asp His Val Thr
 325 330 335
 Arg Val Thr Phe Trp Gly Leu Asp Asp Gly Met Thr Trp Leu Asn Gly
 340 345 350
 Phe Pro Val Arg Gly Arg Thr Asn His Pro Leu Leu Tyr Asp Arg Gln
 355 360 365
 Leu Asn Ala Lys Pro Ala Phe His Ala Leu Val Lys Leu Gly Gln Glu
 370 375 380
 Glu Arg Pro Glu Ala Ala Lys Val Glu Val Gln Lys Ile Glu Ala Lys
 385 390 395 400
 Lys Glu Glu Ala Lys
 405

<210> 133
 <211> 1011
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 133
 atgaaaaata atcaatttag gaaaatccct tccctacata aggtatataa gagtcatttt 60
 ttaattgggg cagctgtaaa tccacttaca cttcaaacac aacaggaact aatcaaaaag 120
 cactttaata gtattacggc agaaaatgaa atgaaatttg aagagttgca acctgagcct 180
 ggacatttta cttttgatgt aggagataaa atgggtcgctt tcgcaaaaga aaatgggtatg 240
 aaagtttagag gtcatacatt aatctggcac aatcaaacac ctgattggat gtttaagaat 300
 gaagatgggt ctgtcacaga tcgagataca cttcttgaaa gaatgaaatt acatattaca 360
 actgtttatgg agcattataa ggggcaaatt tattgttggg atgttgtcaa tgaagcgatt 420
 gctgatgaag gatcagagtt attacgtcac tctaaatgga ctgaaattat tggcgacgat 480
 tttattgaaa aggcatttga gtatgcacat gaagcagacc cagaagcttt actattctat 540
 aatgactata atgagtccca ccctcataag cgagataaaa ttacacact aataaaaaga 600
 ttggtagaca aaggcatacc tattcacggg gttggcttgc aagcacattg gaatttaaca 660
 gacccttctt atgaggagat tagggctgca attgaaaaat atgcctcatt aggcttgga 720
 atacatctta cagaaatgga tgtttcagt ttcaattttg aagatcgaag aacagactta 780
 acagagccga ctaatgaaat gaagactctt caagtagaac gttatacggg atttttcaaa 840
 atacttagag aatatagcca tgtgattagc tctgtcactt ttgggggagc tgcagatgat 900
 tatacttggt tggatgggtt tccagttaga ggaaggaaaa actggccatt tgtttttgac 960
 gaaaaccacc aaccgaaaga atctttctgg ggaattgtcg attttgaata a 1011

<210> 134
 <211> 336
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 134
 Met Lys Asn Asn Gln Phe Arg Lys Ile Pro Ser Leu His Lys Val Tyr
 1 5 10 15
 Lys Ser His Phe Leu Ile Gly Ala Ala Val Asn Pro Leu Thr Leu Gln
 20 25 30
 Thr Gln Gln Glu Leu Ile Lys Lys His Phe Asn Ser Ile Thr Ala Glu
 35 40 45
 Asn Glu Met Lys Phe Glu Glu Leu Gln Pro Glu Pro Gly His Phe Thr
 50 55 60
 Phe Asp Val Gly Asp Lys Met Val Ala Phe Ala Lys Glu Asn Gly Met
 65 70 75 80
 Lys Val Arg Gly His Thr Leu Ile Trp His Asn Gln Thr Pro Asp Trp
 85 90 95
 Met Phe Lys Asn Glu Asp Gly Ser Val Thr Asp Arg Asp Thr Leu Leu
 100 105 110
 Glu Arg Met Lys Leu His Ile Thr Thr Val Met Glu His Tyr Lys Gly

115 120 125
 Gln Ile Tyr Cys Trp Asp Val Asn Glu Ala Ile Ala Asp Glu Gly
 130 135 140
 Ser Glu Leu Leu Arg His Ser Lys Trp Thr Glu Ile Ile Gly Asp Asp
 145 150 155 160
 Phe Ile Glu Lys Ala Phe Glu Tyr Ala His Glu Ala Asp Pro Glu Ala
 165 170 175
 Leu Leu Phe Tyr Asn Asp Tyr Asn Glu Ser His Pro His Lys Arg Asp
 180 185 190
 Lys Ile Tyr Thr Leu Ile Lys Arg Leu Val Asp Lys Gly Ile Pro Ile
 195 200 205
 His Gly Val Gly Leu Gln Ala His Trp Asn Leu Thr Asp Pro Ser Tyr
 210 215 220
 Glu Glu Ile Arg Ala Ala Ile Glu Lys Tyr Ala Ser Leu Gly Leu Glu
 225 230 235 240
 Ile His Leu Thr Glu Met Asp Val Ser Val Phe Asn Phe Glu Asp Arg
 245 250 255
 Arg Thr Asp Leu Thr Glu Pro Thr Asn Glu Met Lys Thr Leu Gln Val
 260 265 270
 Glu Arg Tyr Thr Glu Phe Phe Lys Ile Leu Arg Glu Tyr Ser His Val
 275 280 285
 Ile Ser Ser Val Thr Phe Trp Gly Ala Ala Asp Asp Tyr Thr Trp Leu
 290 295 300
 Asp Gly Phe Pro Val Arg Gly Arg Lys Asn Trp Pro Phe Val Phe Asp
 305 310 315 320
 Glu Asn His Gln Pro Lys Glu Ser Phe Trp Gly Ile Val Asp Phe Glu
 325 330 335

<210> 135
 <211> 1170
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 135
 atgcgacgcc tcacgcgcct tgtcctatat ataggaaccg ccgcgagcgg gacctccgtg 60
 gagaccgttg cgccgaatc gaaacagccg aaagctagcc taaagaatgc gttcgcagac 120
 gattttcgtg tcggcgctgc aattggcacc aatcagggtca tgggcgagga gccaaaatcg 180
 ctgcagggttg tcgcccagca gttcaacaca atcacgcctg agaatctcct caaatgggct 240
 gaggtccacc cagaagcaga ccgctacaac ttcgaaccgt ccgatcgctt cgtcgaattt 300
 ggcgaaaaga acaacatggt catcgtcggc cacacgctcg tgtggcataa ccaaacgccg 360
 gactgggcct ttgagggcaa ggacggcaag ccgctcgatc gcgaaacagc gctcgcccga 420
 atcaaggaaac acattgaaac cgtggtcggc cgatatcgcg gccgatcca tgcttgggac 480
 gtcgtgaacg aggcaatcga cgacaacggc aaacttcgta gtgggcccgt cggagtggcc 540
 ggtcagcgcg gcgaaccgtg gcacgcccgc atcggagacg actacatcca gaaggcgtt 600
 gaattcgcgc acaccgccga ccccgacgct gaactctatt acaacgacta caacgaatgg 660
 caccgaaaaa agatcgaagc catctcgag ctgggtcggt cgctcaaaga gaaggcggt 720
 cgtatcgatg gcctcgggtc ccagggccat tgggggatgg attaccgaa agtcgaagag 780
 atcgatcaca tgctaaccga gtatggcaag ctcggcgtga agctcatgat taccgaactc 840
 gacatcaaca tgcttcgcga gcccgaaccg agtcaacgcg gcgccgatat cactcgcaac 900
 tacgagctca gaaaggagct cgtaccgtat tccgacggac tcccggccga tatgcaaaag 960
 gcactcgcgg cgcgttatgc tgaaatcttc gaagtcttcg ctaagcatcg cgataagctc 1020
 gaccgcgtca cattttgggg cgttcacgac ggccattcat ggctcaacaa ctggcctgtt 1080
 cccggtcgca ctgcctaccc gcttctcttc gacacgaagc ttcagcccaa gccggcattt 1140
 gatgccgtca tcggagtcgc agagcaatga 1170

<210> 136
 <211> 389
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(25)

<400> 136

```

Met Arg Arg Leu Ile Ala Leu Val Leu Tyr Ile Gly Thr Ala Ala Ser
1      5      10      15
Gly Thr Ser Val Glu Thr Val Ala Ala Glu Ser Lys Gln Pro Lys Ala
20      25      30
Ser Leu Lys Asn Ala Phe Ala Asp Asp Phe Arg Val Gly Ala Ala Ile
35      40      45
Gly Thr Asn Gln Val Met Gly Glu Glu Pro Lys Ser Leu Glu Val Val
50      55      60
Ala Gln Gln Phe Asn Thr Ile Thr Pro Glu Asn Leu Leu Lys Trp Ala
65      70      75      80
Glu Val His Pro Glu Ala Asp Arg Tyr Asn Phe Glu Pro Ser Asp Arg
85      90      95
Phe Val Glu Phe Gly Glu Lys Asn Asn Met Phe Ile Val Gly His Thr
100      105      110
Leu Val Trp His Asn Gln Thr Pro Asp Trp Ala Phe Glu Gly Lys Asp
115      120      125
Gly Lys Pro Leu Asp Arg Glu Thr Ala Leu Ala Arg Ile Lys Glu His
130      135      140
Ile Glu Thr Val Val Gly Arg Tyr Arg Gly Arg Ile His Ala Trp Asp
145      150      155      160
Val Val Asn Glu Ala Ile Asp Asp Asn Gly Lys Leu Arg Ser Gly Pro
165      170      175
Val Gly Val Pro Gly Gln Arg Gly Glu Pro Trp His Ala Ala Ile Gly
180      185      190
Asp Asp Tyr Ile Gln Lys Ala Phe Glu Phe Ala His Thr Ala Asp Pro
195      200      205
Asp Ala Glu Leu Tyr Tyr Asn Asp Tyr Asn Glu Trp His Pro Lys Lys
210      215      220
Ile Glu Ala Ile Ser Gln Leu Val Arg Ser Leu Lys Glu Lys Gly Val
225      230      235      240
Arg Ile Asp Gly Leu Gly Leu Gln Gly His Trp Gly Met Asp Tyr Pro
245      250      255
Lys Val Glu Glu Ile Asp His Met Leu Thr Glu Tyr Gly Lys Leu Gly
260      265      270
Val Lys Leu Met Ile Thr Glu Leu Asp Ile Asn Met Leu Pro Gln Pro
275      280      285
Asp Pro Ser Gln Arg Gly Ala Asp Ile Thr Arg Asn Tyr Glu Leu Arg
290      295      300
Lys Glu Leu Asp Pro Tyr Ser Asp Gly Leu Pro Pro Asp Met Gln Lys
305      310      315      320
Ala Leu Ala Ala Arg Tyr Ala Glu Ile Phe Glu Val Phe Ala Lys His
325      330      335
Arg Asp Lys Leu Asp Arg Val Thr Phe Trp Gly Val His Asp Gly His
340      345      350
Ser Trp Leu Asn Asn Trp Pro Val Pro Gly Arg Thr Ala Tyr Pro Leu
355      360      365
Leu Phe Asp Thr Lys Leu Gln Pro Lys Pro Ala Phe Asp Ala Val Ile
370      375      380
Gly Val Ala Glu Gln
385

```

<210> 137

<211> 1044

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 137

```

gtggatcctt cgctgaagga agcagcttcg ggcaagtctt tggatgggggt agcgttgaat      60
gtacgtcagg cagcaggtca ggatacttgc gcctcgaaag tggtaaaacg tcattttaat      120
tccattgtgg ccgagaattg catgaaatgc gaagtgtatt atccggagga agaccatttt      180
gattttacgg aagcggaccg gttggttcgt tttggcgagg agaacgatat ggctgttatc      240
gggcattgcc ttatctggca ttcacagctg gcaccttggt tctgtgtgga caaacaagga      300
aaaacagtaa gtgccgacat cttgaaggag cgtataaaaa aacatatcca gactattgtg      360
acgcactata aagggcgtat aaaggcgtgg gatgtgttga atgaagccat tgaatcggac      420
ggctcctggc gtaaattctc tttttacgag atattaggcg aagagtacat cccgcttatt      480

```

tttcagtatg	ctcatgaggc	agatccggaa	gccgaacttt	actataatga	ttatggcatg	540
gacgggaagg	ctaagcgtga	caaagtagtc	gaattggtaa	agatgctgaa	agatcgtgga	600
ctgcgcatcg	acgcggtagg	tatgcaggga	cacatgggaa	tggattatcc	gtcagtgtcc	660
gaatttgaag	ccagtatact	ggcatttgca	gctgccggag	taaaggtgat	ggtaaccgaa	720
tgggatatga	gtgcattgcc	cacgacacgg	atgggagcca	atatttcgga	cacggtgtct	780
tataaacaat	ccctgaatcc	ctatcccggac	ggtttgcccg	actctgtgtc	tgtggcatgg	840
aataaccgga	tgaaggaatt	tttcggtctt	ttcctgaaac	attcgaatat	cattaccctg	900
gtgacggcgt	gggggggtgac	ggacgggtgac	tcatgggaaga	ataatttccc	tgtgcccgga	960
cgtgtggatt	atcctttatt	gttcgaccgt	gattgccggc	cgaaaccttt	tgtggaagaa	1020
ctgattggaa	aacagaacat	ttaa				1044

<210> 138
 <211> 347
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 138

Val	Asp	Pro	Ser	Leu	Lys	Glu	Ala	Ala	Ser	Gly	Lys	Phe	Leu	Met	Gly
1				5					10					15	
Val	Ala	Leu	Asn	Val	Arg	Gln	Ala	Ala	Gly	Gln	Asp	Thr	Cys	Ala	Ser
			20					25					30		
Lys	Val	Val	Lys	Arg	His	Phe	Asn	Ser	Ile	Val	Ala	Glu	Asn	Cys	Met
		35					40					45			
Lys	Cys	Glu	Val	Ile	His	Pro	Glu	Glu	Asp	His	Phe	Asp	Phe	Thr	Glu
	50					55					60				
Ala	Asp	Arg	Leu	Val	Arg	Phe	Gly	Glu	Glu	Asn	Asp	Met	Ala	Val	Ile
65					70					75				80	
Gly	His	Cys	Leu	Ile	Trp	His	Ser	Gln	Leu	Ala	Pro	Trp	Phe	Cys	Val
			85						90					95	
Asp	Lys	Gln	Gly	Lys	Thr	Val	Ser	Ala	Asp	Ile	Leu	Lys	Glu	Arg	Ile
			100					105					110		
Lys	Lys	His	Ile	Gln	Thr	Ile	Val	Thr	His	Tyr	Lys	Gly	Arg	Ile	Lys
		115						120				125			
Gly	Trp	Asp	Val	Leu	Asn	Glu	Ala	Ile	Glu	Ser	Asp	Gly	Ser	Trp	Arg
	130					135					140				
Lys	Ser	Pro	Phe	Tyr	Glu	Ile	Leu	Gly	Glu	Glu	Tyr	Ile	Pro	Leu	Ile
145					150					155				160	
Phe	Gln	Tyr	Ala	His	Glu	Ala	Asp	Pro	Glu	Ala	Glu	Leu	Tyr	Tyr	Asn
			165					170						175	
Asp	Tyr	Gly	Met	Asp	Gly	Lys	Ala	Lys	Arg	Asp	Lys	Val	Val	Glu	Leu
			180					185					190		
Val	Lys	Met	Leu	Lys	Asp	Arg	Gly	Leu	Arg	Ile	Asp	Ala	Val	Gly	Met
		195					200					205			
Gln	Gly	His	Met	Gly	Met	Asp	Tyr	Pro	Ser	Val	Ser	Glu	Phe	Glu	Ala
	210					215					220				
Ser	Ile	Leu	Ala	Phe	Ala	Ala	Gly	Val	Lys	Val	Met	Val	Thr	Glu	
225					230				235					240	
Trp	Asp	Met	Ser	Ala	Leu	Pro	Thr	Thr	Arg	Met	Gly	Ala	Asn	Ile	Ser
			245						250					255	
Asp	Thr	Val	Ser	Tyr	Lys	Gln	Ser	Leu	Asn	Pro	Tyr	Pro	Asp	Gly	Leu
			260					265					270		
Pro	Asp	Ser	Val	Ser	Val	Ala	Trp	Asn	Asn	Arg	Met	Lys	Glu	Phe	Phe
		275					280					285			
Gly	Leu	Phe	Leu	Lys	His	Ser	Asn	Ile	Ile	Thr	Arg	Val	Thr	Ala	Trp
	290					295					300				
Gly	Val	Thr	Asp	Gly	Asp	Ser	Trp	Lys	Asn	Asn	Phe	Pro	Val	Pro	Gly
305					310					315					320
Arg	Val	Asp	Tyr	Pro	Leu	Leu	Phe	Asp	Arg	Asp	Cys	Arg	Pro	Lys	Pro
			325						330					335	
Phe	Val	Glu	Glu	Leu	Ile	Gly	Lys	Gln	Asn	Ile					
			340					345							

<210> 139
 <211> 1143
 <212> DNA
 <213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 139

atgaaaaaaa	cgattgcaca	tttcacctta	tggatagtgt	tttttctctt	cacttcctgt	60
actgttacgg	cgcagaagaa	tgctaagaat	gcaagagtaa	aaccactac	cctaaaagag	120
gcttaccaag	gtaaattcta	tatcggtact	gcgatgaact	tgagacagat	tcacggagat	180
gatccccaat	ctgaaaatat	tatcaaaaaa	cagttcaatt	ccatagtgtc	cgaaaactgc	240
atgaagagta	tgtatcttca	gccggaggaa	ggaaaatttt	tcttcgatga	tgcggacaag	300
tttgtggatt	ttgggtcttca	gaacaatatg	ttcattatcg	ggcattgtct	gatttggcat	360
tcgcaggcgc	caaaatgggt	tttcaccgac	gaaaatggaa	acacggtttc	tccagaagtt	420
cttaaacaaa	ggatgaaagc	ccatatcacc	gctgtcgttt	cccgtacaa	agggaaaatc	480
aaaggttggg	atgtggtgaa	cgaagccatt	atggaagatg	gttcttaccg	caaaagcaaa	540
ttttacgaga	ttttgggaga	agaatttatt	ccgttggcat	ttcagtatgc	gcatgaagca	600
gacacctgat	cagaacttta	ttacaacgat	tataacgaat	ggtatcccgg	gaaaagagct	660
atggttgacca	aaataatccg	cgatttcaaa	actagaggaa	tccgcatcga	tgccatcgga	720
atgcaggctc	atttcgggat	ggattcgccc	actgtagaag	agtatgaaca	aactattcag	780
ggctatataa	aagaaggcgt	gaaagtcaat	attacggaac	tcgatttaag	tccgcttcct	840
tctccttggg	gaacttccgc	caacgttgct	gatacgagc	agtatcagga	aaaaatgaat	900
ccttacacca	aaggacttcc	tgtcgatgta	gaaaaagcat	gggaaaaccg	ttatctcgat	960
tttttcaaac	ttttcctaag	atatcatcag	catattgagc	gtgtaacttt	ttggggagtg	1020
agcgacatcg	atttcctggaa	aaacgatttt	ccgataagag	gacgtaccga	ttatccacta	1080
ccgtttaacc	gtcaatatca	ggcaaaacct	ttggttcaga	aattaataga	cttaacgaaa	1140
tag						1143

<210> 140

<211> 380

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(24)

<400> 140

Met	Lys	Lys	Thr	Ile	Ala	His	Phe	Thr	Leu	Trp	Ile	Val	Phe	Phe	Leu
1				5					10				15		
Phe	Thr	Ser	Cys	Thr	Val	Thr	Ala	Gln	Lys	Asn	Ala	Lys	Asn	Ala	Arg
			20					25					30		
Val	Lys	Pro	Thr	Thr	Leu	Lys	Glu	Ala	Tyr	Gln	Gly	Lys	Phe	Tyr	Ile
		35					40					45			
Gly	Thr	Ala	Met	Asn	Leu	Arg	Gln	Ile	His	Gly	Asp	Asp	Pro	Gln	Ser
	50					55					60				
Glu	Asn	Ile	Ile	Lys	Lys	Gln	Phe	Asn	Ser	Ile	Val	Ala	Glu	Asn	Cys
	65				70				75						80
Met	Lys	Ser	Met	Tyr	Leu	Gln	Pro	Glu	Glu	Gly	Lys	Phe	Phe	Phe	Asp
				85				90						95	
Asp	Ala	Asp	Lys	Phe	Val	Asp	Phe	Gly	Leu	Gln	Asn	Asn	Met	Phe	Ile
			100					105					110		
Ile	Gly	His	Cys	Leu	Ile	Trp	His	Ser	Gln	Ala	Pro	Lys	Trp	Phe	Phe
		115				120						125			
Thr	Asp	Glu	Asn	Gly	Asn	Thr	Val	Ser	Pro	Glu	Val	Leu	Lys	Gln	Arg
	130					135					140				
Met	Lys	Ala	His	Ile	Thr	Ala	Val	Val	Ser	Arg	Tyr	Lys	Gly	Lys	Ile
	145				150					155					160
Lys	Gly	Trp	Asp	Val	Val	Asn	Glu	Ala	Ile	Met	Glu	Asp	Gly	Ser	Tyr
			165					170						175	
Arg	Lys	Ser	Lys	Phe	Tyr	Glu	Ile	Leu	Gly	Glu	Glu	Phe	Ile	Pro	Leu
			180					185					190		
Ala	Phe	Gln	Tyr	Ala	His	Glu	Ala	Asp	Pro	Asp	Ala	Glu	Leu	Tyr	Tyr
		195					200					205			
Asn	Asp	Tyr	Asn	Glu	Trp	Tyr	Pro	Gly	Lys	Arg	Ala	Met	Val	Thr	Lys
	210					215					220				
Ile	Ile	Arg	Asp	Phe	Lys	Thr	Arg	Gly	Ile	Arg	Ile	Asp	Ala	Ile	Gly
	225				230				235						240
Met	Gln	Ala	His	Phe	Gly	Met	Asp	Ser	Pro	Thr	Val	Glu	Glu	Tyr	Glu

245 250 255
 Gln Thr Ile Gln Gly Tyr Ile Lys Glu Gly Val Lys Val Asn Ile Thr
 260 265 270
 Glu Leu Asp Leu Ser Pro Leu Pro Ser Pro Trp Gly Thr Ser Ala Asn
 275 280 285
 Val Ala Asp Thr Gln Gln Tyr Gln Glu Lys Met Asn Pro Tyr Thr Lys
 290 295 300
 Gly Leu Pro Val Asp Val Glu Lys Ala Trp Glu Asn Arg Tyr Leu Asp
 305 310 315 320
 Phe Phe Lys Leu Phe Leu Lys Tyr His Gln His Ile Glu Arg Val Thr
 325 330 335
 Phe Trp Gly Val Ser Asp Ile Asp Ser Trp Lys Asn Asp Phe Pro Ile
 340 345 350
 Arg Gly Arg Thr Asp Tyr Pro Leu Pro Phe Asn Arg Gln Tyr Gln Ala
 355 360 365
 Lys Pro Leu Val Gln Lys Leu Ile Asp Leu Thr Lys
 370 375 380

<210> 141
 <211> 1134
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 141
 atgaatatct cacgcagaca actactggcg ctacggggtg ctacggcggc gatcacagca 60
 gccaaattac aggcggcaga aaaagccagc gccgcgaccg gcttgcgcga tgcctacaaa 120
 aatgattttt tgattggcgc tgcgctgagt gcatcgatca ttcaacagca agatccacag 180
 ctagtgtcac tgattaataa agactttaat tccatcaccc cagaaaactg tatgaaatgg 240
 ggcgagatgc gcaatgatga cggcagctgg aagtggcagg atgcagacgc atttgtcgag 300
 tatggaagca aatacaaaact acatatgggtc ggccacacat tgggggtggca cagccagatt 360
 cccgatagcg tgtttaaaaa taaagacggg agctatatatt ccaaaaccga actcgcgaaa 420
 aaacaaaaag aacacatcac cactattggt ggccgctaca aaggcaaaact tgccgcgtgg 480
 gatgtggtga atgaagctgt cggcgatgac aacaaaatgc gcgatagtca ctggtataaa 540
 atcatgggcg atgatitttct cgttaatgca tttaaccttg ctcatgaagt agatccgaag 600
 gcgcatctga tgtacaacga ctacaacaac gagcgcccgg aaaaacgcca ggcgactatc 660
 gatattgata agcgtctgca acaacgcggg acaccaatcc atggtttggg catgcaagcg 720
 catatcggat tggaaaacaa tatgcaggat ttggaagata gtattctcgc ctattcagca 780
 ttgggtttta aaatccatct caccgaacta gatatagatg tgctgccctc tgtatggaat 840
 ttaccggtgg ccgaaatttc taccgcgttt gaatacaagc cggaaacgca tccttatata 900
 aaaggtttgc cgaaagagat tgatgaaaaa cttgcaaaag cctatgaatc gctattttaa 960
 atattgctta aacatcgca caaaatagat agagttacgt tttggggcgt aagcgatgat 1020
 gccagctggc tcaatgatgt cccaatcaat ggcaaacca actatccgtt attgtttaac 1080
 cgtcaacgcc aacctaaagc tgcttatttc cgtttgctgg atttaaaacg ctac 1134

<210> 142
 <211> 377
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(25)

<400> 142
 Met Asn Ile Ser Arg Arg Gln Leu Leu Ala Leu Thr Gly Ala Thr Ala
 1 5 10 15
 Ala Ile Thr Ala Lys Leu Gln Ala Glu Lys Ala Ser Ala Ala
 20 25 30
 Thr Gly Leu Arg Asp Ala Tyr Lys Asn Asp Phe Leu Ile Gly Ala Ala
 35 40 45
 Leu Ser Ala Ser Ile Ile Gln Gln Asp Pro Gln Leu Val Ala Leu
 50 55 60
 Ile Asn Lys Asp Phe Asn Ser Ile Thr Pro Glu Asn Cys Met Lys Trp
 65 70 75 80

Gly Glu Met Arg Asn Asp Asp Gly Ser Trp Lys Trp Gln Asp Ala Asp
 85 95
 Ala Phe Val Glu Tyr Gly Ser Lys Tyr Lys Leu His Met Val Gly His
 100 105 110
 Thr Leu Gly Trp His Ser Gln Ile Pro Asp Ser Val Phe Lys Asn Lys
 115 120 125
 Asp Gly Ser Tyr Ile Ser Lys Thr Glu Leu Ala Lys Lys Gln Lys Glu
 130 135 140
 His Ile Thr Thr Ile Val Gly Arg Tyr Lys Gly Lys Leu Ala Ala Trp
 145 150 155 160
 Asp Val Val Asn Glu Ala Val Gly Asp Asp Asn Lys Met Arg Asp Ser
 165 170 175
 His Trp Tyr Lys Ile Met Gly Asp Asp Phe Leu Val Asn Ala Phe Asn
 180 185 190
 Leu Ala His Glu Val Asp Pro Lys Ala His Leu Met Tyr Asn Asp Tyr
 195 200 205
 Asn Asn Glu Arg Pro Glu Lys Arg Gln Ala Thr Ile Asp Met Ile Lys
 210 215 220
 Arg Leu Gln Gln Arg Gly Thr Pro Ile His Gly Leu Gly Met Gln Ala
 225 230 235 240
 His Ile Gly Leu Glu Thr Asn Met Gln Asp Phe Glu Asp Ser Ile Leu
 245 250 255
 Ala Tyr Ser Ala Leu Gly Leu Lys Ile His Leu Thr Glu Leu Asp Ile
 260 265 270
 Asp Val Leu Pro Ser Val Trp Asn Leu Pro Val Ala Glu Ile Ser Thr
 275 285
 Arg Phe Glu Tyr Lys Pro Glu Arg Asp Pro Tyr Thr Lys Gly Leu Pro
 290 295 300
 Lys Glu Ile Asp Glu Lys Leu Ala Lys Ala Tyr Glu Ser Leu Phe Lys
 305 310 315 320
 Ile Leu Leu Lys His Arg Asp Lys Ile Asp Arg Val Thr Phe Trp Gly
 325 330 335
 Val Ser Asp Asp Ala Ser Trp Leu Asn Asp Phe Pro Ile Asn Gly Arg
 340 345 350
 Thr Asn Tyr Pro Leu Leu Phe Asn Arg Gln Arg Gln Pro Lys Ala Ala
 355 360 365
 Tyr Phe Arg Leu Leu Asp Leu Lys Arg
 370 375

<210> 143
 <211> 3285
 <212> DNA
 <213> Bacteria

<400> 143
 atgagtttaa aaataaataa aatcatatca tttatcatag ttttttcgat gggtttttggg 60
 acgttaatgt atgtgccaca tctaaaagca tttgcggata ataccggtat taattttgggt 120
 tctaattggtg attttgaatc aggcacaaatt gatggctggt ttaaaCaagg taatccgaca 180
 ttaacagtaa caactgagca ggcaattggg caatacagta tgaaagttac aggtagaaca 240
 cagacatatg aaggaccgcg atatatgcttt ttggggaaaa tgcagaaagg tgaatcatat 300
 aacgtatcac ttaaagtttag acttgtttct ggacaaaatt catctaattcc ttgatcact 360
 gtaactatgt ttagagaaga tgacaatggc aatcattatg acacaatagt ttggcaaaaa 420
 caagtttctg aagattcatg gactactgta agtgggactt atacattgga ttatactgga 480
 acattaaaaa cattatatat gtatgtagaa tcaccggtc caacgcttga atatttatatt 540
 gatgatgttg tagtcacacc gcaaaatcca acgcaaata gaaatgtagt tgccaatgga 600
 acttttgaat atgaaaatac ttctggatgg gttggaacag gttcatctgt tgttaaagca 660
 gtatatgggtg atgctcacag cggagattat agcttattga cgacaggaag gacagctaac 720
 tggaaatggc ctagttagta ttgactggc aaaatagttc cgggacaaca atacaatgtg 780
 gatttttggg taaaatttat tgatggcaat gatacagagc aaatcaaggc tactgttaaa 840
 gcgacttctg acaaagacaa ttatatacaa gttaatgatt ttgcagatgt aagtaaagg 900
 gaatggacag aaataaaagg cagttttact ttacctgttg cagattacag cggcattagc 960
 atctatgtgg aatctcaaaa tcctacttta gagttttaca ttgatgattt ttctgtaata 1020
 ggtgaaattg caaataatca gattactatt caaaatgaca ttccagattt gtactctgta 1080
 tttaaagatt attttcctat aggcgtttgc gttgatccaa gtagattaaa tgatactgat 1140
 ccgcatgctc aattgacggc taaacatttt aatatgcttg ttgcagaaaa cgccatgaaa 1200
 cctgaaagtt tacaaccac agaaggaaat tttacttttg ataattgctga taagattggt 1260
 gattatgcaa tatcacataa tatgaagatg agggggcata ctttactttg gcataatcaa 1320
 gttccagatt gggttttcca agatccgtct gacccatcca agcctgcttc gagagattta 1380
 ctattacaaa gattaaaaac tcatattaca actgtgttag accattttta aacaaagtat 1440

```

ggttctcaga atccaataat tggatgggat gtcgtaaagt aagttcttga tgataatggc 1500
agtttgagaa attcgaagtg gttgcaaatt attggacccg actatataga aaaagccttt 1560
gaatatgcac atgaagcggg tccatcgatg aagttgttta ttaatgatta caatatcgaa 1620
aataatggcg ttaaaactca agctatgtat gacttggtta aaagattaaa gagtgaaggc 1680
gttcctatag atggaatagg gatgcaaatg cacataaata taaattccaa tatagataat 1740
ataaaagcat caatagaaaa actggcatcg ttaggcgttg aaatacaagt aactgaatta 1800
gatatgaaca tgaacggtaa tatatctaac gaagcattgc tcaagcaagc tagattgtat 1860
aaacaattat ttgacttatt taaagcagag aaacaatata taactgctgt agttttttgg 1920
ggagtttcag acgatgtaac ttggccttagc aagccaaatg ctccgctact ttttgatata 1980
aagttgcaag caaagccagc atactgggca atagtagatc cgaataaagc tacaccagac 2040
attcaatctg caaaggcttt ggaaggatca ccgacaatgg gtacaaatgt tgataactct 2100
tggaactctg taaagccggt atatgcaaat actttttagg aagggctcgg cggagcaact 2160
gctgctgtta agtctatgtg ggatactaaa aacttgtatt tgtagtaca agtttcagac 2220
aataccccat ctagtaatga tggatttgag attttttagg ataagaatga tgacaaatcc 2280
acttcttatg aaactgatga tgaacattat acaattaaga gggatggtac agggagtcca 2340
gatattacca aatatgtgac ttctaattgct gacggatatg tagcacagct agctattcca 2400
attgaagata ttaatcctgc acttaatgat aaaattggat ttgacattag aataaatgat 2460
gataaaggta ttggtaatat agatgcaata acagtttggg acgattatac aaacagtcaa 2520
gatacataata catcgtattt tggcgattta gtattatcaa aacctgcaca aattgcaaca 2580
gctatatatg gcactcctgt tattgatggg aaagtagatg atatttggaa taatgttgaa 2640
gctatttcaa caaatatcgt ggttttgggt tcaaatgggt ctactgcgac agcaaaaatg 2700
atgtgggacg ataagtacct ttatgttttg gcggatgtta cagattcaaa tctgaacaaa 2760
tctagtgtta atccatatga acaagattct gtagaagttt ttgtagatca aaataatgac 2820
aagacgacat attataaaaa tgatgatgga cagtttagag ttaactatga caatgaacaa 2880
agcttttggg gaagcactaa ttcaaatgga tttaaatcgg caactagtct tacacaaagt 2940
ggatatattg tagaagaagc tattccttgg acgagtatca ctccatcaaa tggcactatc 3000
ataggatttg acttgcaagt taatgatgca gatgaaaatg gtaagaggac aggtattgta 3060
acatgggtgtg atccgagcgg aaattcatgg caagatactt ctgggttttg gaatttattg 3120
cttacaggta aaccatccgg tgttgttaca aaaagaatgg cgtttaacga cataaaagac 3180
agttgggcaa aagatgcaat agaagtatta gcatcaaggc acatagtaga aggtatgaca 3240
gacactcagt atgaaccaa caagacagta acgagagcgg aataa 3285

```

<210> 144

<211> 1094

<212> PRT

<213> Bacteria

<220>

<221> SIGNAL

<222> (1)...(32)

<400> 144

```

Met Ser Leu Lys Ile Asn Lys Ile Ile Ser Phe Ile Ile Val Phe Ser
1      5      10      15
Met Val Phe Gly Thr Leu Met Tyr Val Pro His Leu Lys Ala Phe Ala
20      25      30
Asp Asn Thr Gly Ile Asn Leu Val Ser Asn Gly Asp Phe Glu Ser Gly
35      40      45
Thr Ile Asp Gly Trp Phe Lys Gln Gly Asn Pro Thr Leu Thr Val Thr
50      55      60
Thr Glu Gln Ala Ile Gly Gln Tyr Ser Met Lys Val Thr Gly Arg Thr
65      70      75      80
Gln Thr Tyr Glu Gly Pro Ala Tyr Ser Phe Leu Gly Lys Met Gln Lys
85      90      95
Gly Glu Ser Tyr Asn Val Ser Leu Lys Val Arg Leu Val Ser Gly Gln
100     105     110
Asn Ser Ser Asn Pro Leu Ile Thr Val Thr Met Phe Arg Glu Asp Asp
115     120     125
Asn Gly Asn His Tyr Asp Thr Ile Val Trp Gln Lys Gln Val Ser Glu
130     135     140
Asp Ser Trp Thr Thr Val Ser Gly Thr Tyr Thr Leu Asp Tyr Thr Gly
145     150     155     160
Thr Leu Lys Thr Leu Tyr Met Tyr Val Glu Ser Pro Asp Pro Thr Leu
165     170     175
Glu Tyr Tyr Ile Asp Asp Val Val Val Thr Pro Gln Asn Pro Thr Gln
180     185     190
Ile Gly Asn Val Val Ala Asn Gly Thr Phe Glu Asn Glu Asn Thr Ser
195     200     205
Gly Trp Val Gly Thr Gly Ser Ser Val Val Lys Ala Val Tyr Gly Asp

```

210	215	220
Ala His Ser Gly Asp Tyr	Ser Leu Leu Thr Thr	Gly Arg Thr Ala Asn
225	230	235
Trp Asn Gly Pro Ser Tyr	Asp Leu Thr Gly Lys	Ile Val Pro Gly Gln
245	250	255
Gln Tyr Asn Val Asp Phe Trp Val Lys Phe	Ile Asp Gly Asn Asp Thr	
260	265	270
Glu Gln Ile Lys Ala Thr Val Lys Ala Thr	Ser Asp Lys Asp Asn Tyr	
275	280	285
Ile Gln Val Asn Asp Phe Ala Asp Val Ser Lys	Gly Glu Trp Thr Glu	
290	295	300
Ile Lys Gly Ser Phe Thr Leu Pro Val Ala Asp	Tyr Ser Gly Ile Ser	
305	310	315
Ile Tyr Val Glu Ser Gln Asn Pro Thr Leu Glu	Phe Tyr Ile Asp Asp	
325	330	335
Phe Ser Val Ile Gly Glu Ile Ala Asn Asn	Gln Ile Thr Ile Gln Asn	
340	345	350
Asp Ile Pro Asp Leu Tyr Ser Val Phe Lys	Asp Tyr Phe Pro Ile Gly	
355	360	365
Val Ala Val Asp Pro Ser Arg Leu Asn Asp	Thr Asp Pro His Ala Gln	
370	375	380
Leu Thr Ala Lys His Phe Asn Met Leu Val Ala	Glu Asn Ala Met Lys	
385	390	395
Pro Glu Ser Leu Gln Pro Thr Glu Gly Asn Phe	Thr Phe Asp Asn Ala	
405	410	415
Asp Lys Ile Val Asp Tyr Ala Ile Ser His	Asn Met Lys Met Arg Gly	
420	425	430
His Thr Leu Leu Trp His Asn Gln Val Pro	Asp Trp Phe Phe Gln Asp	
435	440	445
Pro Ser Asp Pro Ser Lys Pro Ala Ser Arg	Asp Leu Leu Leu Gln Arg	
450	455	460
Leu Lys Thr His Ile Thr Thr Val Leu Asp	His Phe Lys Thr Lys Tyr	
465	470	475
Gly Ser Gln Asn Pro Ile Ile Gly Trp Asp	Val Val Asn Glu Val Leu	
485	490	495
Asp Asp Asn Gly Ser Leu Arg Asn Ser Lys	Trp Leu Gln Ile Ile Gly	
500	505	510
Pro Asp Tyr Ile Glu Lys Ala Phe Glu Tyr	Ala His Glu Ala Asp Pro	
515	520	525
Ser Met Lys Leu Phe Ile Asn Asp Tyr Asn	Ile Glu Asn Asn Gly Val	
530	535	540
Lys Thr Gln Ala Met Tyr Asp Leu Val Lys	Arg Leu Lys Ser Glu Gly	
545	550	555
Val Pro Ile Asp Gly Ile Gly Met Gln Met	His Ile Asn Ile Asn Ser	
565	570	575
Asn Ile Asp Asn Ile Lys Ala Ser Ile Glu	Lys Leu Ala Ser Leu Gly	
580	585	590
Val Glu Ile Gln Val Thr Glu Leu Asp Met	Asn Met Asn Gly Asn Ile	
595	600	605
Ser Asn Glu Ala Leu Leu Lys Gln Ala Arg	Leu Tyr Lys Gln Leu Phe	
610	615	620
Asp Leu Phe Lys Ala Glu Lys Gln Tyr Ile	Thr Ala Val Val Phe Trp	
625	630	635
Gly Val Ser Asp Asp Val Thr Trp Leu Ser	Lys Pro Asn Ala Pro Leu	
645	650	655
Leu Phe Asp Thr Lys Leu Gln Ala Lys Pro	Ala Tyr Trp Ala Ile Val	
660	665	670
Asp Pro Asn Lys Ala Thr Pro Asp Ile Gln	Ser Ala Lys Ala Leu Glu	
675	680	685
Gly Ser Pro Thr Met Gly Thr Asn Val Asp	Asn Ser Trp Lys Leu Val	
690	695	700
Lys Pro Leu Tyr Ala Asn Thr Phe Val Glu	Gly Ser Val Gly Ala Thr	
705	710	715
Ala Ala Val Lys Ser Met Trp Asp Thr Lys	Asn Leu Tyr Leu Leu Val	
725	730	735
Gln Val Ser Asp Asn Thr Pro Ser Ser Asn	Asp Gly Ile Glu Ile Phe	
740	745	750
Val Asp Lys Asn Asp Asp Lys Ser Thr Ser	Tyr Glu Thr Asp Asp Glu	
755	760	765

His Tyr Thr Ile Lys Arg Asp Gly Thr Gly Ser Ser Asp Ile Thr Lys
 770 775 780
 Tyr Val Thr Ser Asn Ala Asp Gly Tyr Val Ala Gln Leu Ala Ile Pro
 785 790 795 800
 Ile Glu Asp Ile Asn Pro Ala Leu Asn Asp Lys Ile Gly Phe Asp Ile
 805 810 815
 Arg Ile Asn Asp Asp Lys Gly Ile Gly Asn Ile Asp Ala Ile Thr Val
 820 825 830
 Trp Asn Asp Tyr Thr Asn Ser Gln Asp Thr Asn Thr Ser Tyr Phe Gly
 835 840 845
 Asp Leu Val Leu Ser Lys Pro Ala Gln Ile Ala Thr Ala Ile Tyr Gly
 850 855 860
 Thr Pro Val Ile Asp Gly Lys Val Asp Asp Ile Trp Asn Asn Val Glu
 865 870 875 880
 Ala Ile Ser Thr Asn Thr Trp Val Leu Gly Ser Asn Gly Ala Thr Ala
 885 890 895
 Thr Ala Lys Met Met Trp Asp Asp Lys Tyr Leu Tyr Val Leu Ala Asp
 900 905 910
 Val Thr Asp Ser Asn Leu Asn Lys Ser Ser Val Asn Pro Tyr Glu Gln
 915 920 925
 Asp Ser Val Glu Val Phe Val Asp Gln Asn Asn Asp Lys Thr Thr Tyr
 930 935 940
 Tyr Lys Asn Asp Asp Gly Gln Phe Arg Val Asn Tyr Asp Asn Glu Gln
 945 950 955 960
 Ser Phe Gly Gly Ser Thr Asn Ser Asn Gly Phe Lys Ser Ala Thr Ser
 965 970 975
 Leu Thr Gln Ser Gly Tyr Ile Val Glu Glu Ala Ile Pro Trp Thr Ser
 980 985 990
 Ile Thr Pro Ser Asn Gly Thr Ile Ile Gly Phe Asp Leu Gln Val Asn
 995 1000 1005
 Asp Ala Asp Glu Asn Gly Lys Arg Thr Gly Ile Val Thr Trp Cys Asp
 1010 1015 1020
 Pro Ser Gly Asn Ser Trp Gln Asp Thr Ser Gly Phe Gly Asn Leu Leu
 1025 1030 1035 1040
 Leu Thr Gly Lys Pro Ser Gly Val Gly Thr Lys Arg Met Ala Phe Asn
 1045 1050 1055
 Asp Ile Lys Asp Ser Trp Ala Lys Asp Ala Ile Glu Val Leu Ala Ser
 1060 1065 1070
 Arg His Ile Val Glu Gly Met Thr Asp Thr Gln Tyr Glu Pro Asn Lys
 1075 1080 1085
 Thr Val Thr Arg Ala Glu
 1090

<210> 145
 <211> 1629
 <212> DNA
 <213> Eukaryote

<400> 145
 atgaagattg tggatacaac ttccgcagag ataaagattg aaatggaacc tgaaaaagag 60
 atacctgtc tgaagaagat actaaaagac tacttcaaag tcggagttgc actgccgtcc 120
 aaggctcttc tcaaccgaa ggacatagaa ctcatacga aacacttcaa cagcatcacc 180
 gcagaaaacg agatgaaacc ggatagtctg ctgcgggca tcgaaaacgg taagtgaag 240
 ttcaggtttg aaacagcaga caaatacatt cagttcgtcg aggaaaacgg catggttata 300
 agaggtcaca cactggtgtg gcacaaccag acaccgact ggttcttcaa agacgaaaac 360
 ggaacacctc tctcaaaga agcgatgacg gaaagactca aagagtacat ccacaccgtt 420
 gtcggacact tcaaaggaaa agtctacgca tgggacgtgg tgaacgaagc ggtcgatccg 480
 aaccagccgg atggactgag aagatccacc tgggtaccaga tcatggggcc tgactacata 540
 gaactcgcct tcaagttcgc aagagaagca gatccagatg caaaactctt ctacaacgac 600
 tacaacacat tcgagcccag aaagagagat atcatctaca acctcgtgaa ggatctcaag 660
 gagaagggag tcatcgatgg gataggcatg cagtgtcaca tcagtcttgc aacagacatc 720
 aaacagatcg aagaggccat caaaaagttc agcaccatac ccggtataga aattcacatc 780
 acagaactcg atatgagtgt ctacagagat tccagttcca actaccaga ggcaccgagg 840
 acggcactca tcgaacaggc tcacaaaatg atgcagctct ttgagatttt caagaagtac 900
 agcaacgtga tcacgaacgt cacattcttg ggtctcaagg acgattactc ctggagagca 960
 acaagaagaa accactggcc gctcatcttc gacaaagatc accaggcgaa actcgcttac 1020
 tgggcgatag tggcacctga ggtccttcca ccacttccaa aagaaagcag gatctccgaa 1080
 ggcgaggctg tggtagtggg gatgatggat gactcgtacc tgatgtcgaa gccgatagag 1140
 atccttgacg aagaagggaa cgtgaaggca acgatcaggg cggtgtggaa agacagcacg 1200


```

atctacatct acggagaggt acaggacaag acgaaaaaac cagcagaaga cggagtggcc 1260
atattcatca acccgaacaa cgaaagaaca ccctatctgc agcctgatga cacctacgct 1320
gtgctgtgga caaactggaa gacggagggtc aacagagaag acgtacaggt gaagaaattc 1380
gttgggcctg gctttagaag atacagcttc gagatgtcga tcacgatacc ggggtgtggag 1440
ttcaagaaag acagctacat aggattcgac gctgcggtga tagacgacgg gaagtggtag 1500
agctggagcg acacgacgaa cagccagaag acgaacacga tgaactacgg aacgctgaaa 1560
ctcgaaggaa taatggtagc gacagcaaaa tacggaacac cggtcatcga tggagagata 1620
gacgagtaa

```

<210> 146

<211> 542

<212> PRT

<213> Eukaryote

<400> 146

```

Met Lys Ile Val Asp Thr Thr Ser Ala Glu Ile Lys Ile Glu Met Glu
1      5      10      15
Pro Glu Lys Glu Ile Pro Ala Leu Lys Glu Val Leu Lys Asp Tyr Phe
20      25      30
Lys Val Gly Val Ala Leu Pro Ser Lys Val Phe Leu Asn Pro Lys Asp
35      40      45
Ile Glu Leu Ile Thr Lys His Phe Asn Ser Ile Thr Ala Glu Asn Glu
50      55      60
Met Lys Pro Asp Ser Leu Leu Ala Gly Ile Glu Asn Gly Lys Leu Lys
65      70      75      80
Phe Arg Phe Glu Thr Ala Asp Lys Tyr Ile Gln Phe Val Glu Glu Asn
85      90      95
Gly Met Val Ile Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro
100     105     110
Asp Trp Phe Phe Lys Asp Glu Asn Gly Asn Leu Leu Ser Lys Glu Ala
115     120     125
Met Thr Glu Arg Leu Lys Glu Tyr Ile His Thr Val Val Gly His Phe
130     135     140
Lys Gly Lys Val Tyr Ala Trp Asp Val Val Asn Glu Ala Val Asp Pro
145     150     155     160
Asn Gln Pro Asp Gly Leu Arg Arg Ser Thr Trp Tyr Gln Ile Met Gly
165     170     175
Pro Asp Tyr Ile Glu Leu Ala Phe Lys Phe Ala Arg Glu Ala Asp Pro
180     185     190
Asp Ala Lys Leu Phe Tyr Asn Asp Tyr Asn Thr Phe Glu Pro Arg Lys
195     200     205
Arg Asp Ile Ile Tyr Asn Leu Val Lys Asp Leu Lys Glu Lys Gly Leu
210     215     220
Ile Asp Gly Ile Gly Met Gln Cys His Ile Ser Leu Ala Thr Asp Ile
225     230     235     240
Lys Gln Ile Glu Glu Ala Ile Lys Lys Phe Ser Thr Ile Pro Gly Ile
245     250     255
Glu Ile His Ile Thr Glu Leu Asp Met Ser Val Tyr Arg Asp Ser Ser
260     265     270
Ser Asn Tyr Pro Glu Ala Pro Arg Thr Ala Leu Ile Glu Gln Ala His
275     280     285
Lys Met Met Gln Leu Phe Glu Ile Phe Lys Lys Tyr Ser Asn Val Ile
290     295     300
Thr Asn Val Thr Phe Trp Gly Leu Lys Asp Asp Tyr Ser Trp Arg Ala
305     310     315     320
Thr Arg Arg Asn Asp Trp Pro Leu Ile Phe Asp Lys Asp His Gln Ala
325     330     335
Lys Leu Ala Tyr Trp Ala Ile Val Ala Pro Glu Val Leu Pro Pro Leu
340     345     350
Pro Lys Glu Ser Arg Ile Ser Glu Gly Glu Ala Val Val Val Gly Met
355     360     365
Met Asp Asp Ser Tyr Leu Met Ser Lys Pro Ile Glu Ile Leu Asp Glu
370     375     380
Glu Gly Asn Val Lys Ala Thr Ile Arg Ala Val Trp Lys Asp Ser Thr
385     390     395     400
Ile Tyr Ile Tyr Gly Glu Val Gln Asp Lys Thr Lys Lys Pro Ala Glu
405     410     415
Asp Gly Val Ala Ile Phe Ile Asn Pro Asn Asn Glu Arg Thr Pro Tyr
420     425     430

```

Leu Gln Pro Asp Asp Thr Tyr Ala Val Leu Trp Thr Asn Trp Lys Thr
 435 440 445
 Glu Val Asn Arg Glu Asp Val Gln Val Lys Lys Phe Val Gly Pro Gly
 450 455 460
 Phe Arg Arg Tyr Ser Phe Glu Met Ser Ile Thr Ile Pro Gly Val Glu
 465 470 475 480
 Phe Lys Lys Asp Ser Tyr Ile Gly Phe Asp Ala Ala Val Ile Asp Asp
 485 490 495
 Gly Lys Trp Tyr Ser Trp Ser Asp Thr Thr Asn Ser Gln Lys Thr Asn
 500 505 510
 Thr Met Asn Tyr Gly Thr Leu Lys Leu Glu Gly Ile Met Val Ala Thr
 515 520 525
 Ala Lys Tyr Gly Thr Pro Val Ile Asp Gly Glu Ile Asp Glu
 530 535 540

<210> 147
 <211> 1146
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 147
 atgactttct ctcgacggca atttttgctg caaacctccg ccggcctggc acttttgagc 60
 actgccaaaa tgcgcgcttt cgcccgcgca gtcgatgaag tgggccttaa agaccatttt 120
 aaagaccatt ttcataattgg gactgccatc agcggtcgac tgatgacgga aatgccggcc 180
 ttttaccgcg acctggttac ccgtgaattc agtgccatta ccatggaaaa cgacatgaaa 240
 tgggagcgct tgcattccaa agaaggccaa tgggattggg agattgccga caaattcgtc 300
 aattttggcg aagaaaacga catgtacatt gtcgggcatg ttcctggtctg gcactcacag 360
 accccggatt ggggtgtcca ggattccaga ggcaagccca tttctcgca cgctttgctg 420
 aaacgcatgc gccaccagat tgaacagatg gcgggcccgt ataagggccg ggtacacgcg 480
 tgggatgtgg tcaatgaggc ggtggacgag gaccaaggct ggcgcaaaaag cccgtggttt 540
 aacattattg ggccccagtt tatggagcac gccttcaatt acgcccacga agtggacccc 600
 gacgctcacc tgtttgtacaa cgactacaat atgcacggtc gggaaaaacg cgaattcgtc 660
 ctggatttca tcaaaagata caagaaaaaa ggcattccga tccagggcat aggcattgcaa 720
 ggccatgtgg gcctgagctt tcccgatatc agcgagtttg agaaaagcct gcaagcctac 780
 gccaaacagg gcatgcggat gcacattacc gagctggata tggacgtgtt accagtggcc 840
 tgggatcaca ttggcgccga gatttccacc gagtttgact acgctgatga actggacccc 900
 tggcccaaaag ggctgccgga agaagtcgaa caggaattta ccgatcgcta caccgctttc 960
 tttaaactgt ttttgaaata ccgcatgat attgaaaggg tcaccttctg gggaaccgga 1020
 gatgcggaat cgtggaaaaa taatttccca gtaagggggc gcaccaacta cccgctgctg 1080
 tttgatcgcc gataccgcag aaaaccggcc tatgattcga ttgtcgaact gaccaaaaac 1140
 ctttaa

<210> 148
 <211> 381
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 148
 Met Thr Phe Ser Arg Arg Gln Phe Leu Leu Gln Thr Ser Ala Gly Leu
 1 5 10 15
 Ala Leu Leu Ser Thr Ala Lys Met Arg Ala Phe Ala Arg Ala Val Asp
 20 25 30
 Glu Val Gly Leu Lys Asp His Phe Lys Asp His Phe His Ile Gly Thr
 35 40 45
 Ala Ile Ser Gly Arg Leu Met Thr Glu Met Pro Ala Phe Tyr Arg Asp
 50 55 60
 Leu Val Thr Arg Glu Phe Ser Ala Ile Thr Met Glu Asn Asp Met Lys
 65 70 75 80
 Trp Glu Arg Leu His Pro Lys Glu Gly Gln Trp Asp Trp Glu Ile Ala
 85 90 95

Asp Lys Phe Val Asn Phe Gly Glu Glu Asn Asp Met Tyr Ile Val Gly
 100 105 110
 His Val Leu Val Trp His Ser Gln Thr Pro Asp Trp Val Phe Gln Asp
 115 120 125
 Ser Arg Gly Lys Pro Ile Ser Arg Asp Ala Leu Leu Lys Arg Met Arg
 130 135 140
 His Gln Ile Glu Gln Met Ala Gly Arg Tyr Lys Gly Arg Val His Ala
 145 150 155 160
 Trp Asp Val Val Asn Glu Ala Val Asp Glu Asp Gln Gly Trp Arg Lys
 165 170 175
 Ser Pro Trp Phe Asn Ile Ile Gly Pro Glu Phe Met Glu His Ala Phe
 180 185 190
 Asn Tyr Ala His Glu Val Asp Pro Asp Ala His Leu Leu Tyr Asn Asp
 195 200 205
 Tyr Asn Met His Gly Arg Glu Lys Arg Glu Phe Val Leu Asp Phe Ile
 210 215 220
 Lys Arg Tyr Lys Lys Lys Gly Ile Pro Ile Gln Gly Ile Gly Met Gln
 225 230 235 240
 Gly His Val Gly Leu Ser Phe Pro Asp Ile Ser Glu Phe Glu Lys Ser
 245 250 255
 Leu Gln Ala Tyr Ala Lys Gln Gly Met Arg Met His Ile Thr Glu Leu
 260 265 270
 Asp Met Asp Val Leu Pro Val Ala Trp Asp His Ile Gly Ala Glu Ile
 275 280 285
 Ser Thr Glu Phe Asp Tyr Ala Asp Glu Leu Asp Pro Trp Pro Lys Gly
 290 295 300
 Leu Pro Glu Glu Val Glu Gln Glu Phe Thr Asp Arg Tyr Thr Ala Phe
 305 310 315 320
 Phe Lys Leu Phe Leu Lys Tyr Arg Asp Asp Ile Glu Arg Val Thr Phe
 325 330 335
 Trp Gly Thr Gly Asp Ala Glu Ser Trp Lys Asn Asn Phe Pro Val Arg
 340 345 350
 Gly Arg Thr Asn Tyr Pro Leu Leu Phe Asp Arg Arg Tyr Arg Arg Lys
 355 360 365
 Pro Ala Tyr Asp Ser Ile Val Glu Leu Thr Lys Asn Leu
 370 375 380

<210> 149
 <211> 1044
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 149
 atgaagaagc tttttgtcgc ggtcgttttg ttgcccttag caactttttt cgcgtcggac 60
 ggattggagg gagaaccttt gagatcggtta gccgagaaac ttggcatcta catcggttac 120
 gcttcgatca accattttctg gactcttccg gattcaaaca agtacacaga agtggcaaag 180
 agggagtcca acatactcac gccagagaac caaatgaagt gggacagcct tcaccagag 240
 cctgacaggt acaacttcac ttacgcagag cgtcatgtcg agttcgtttt ggaaaacaac 300
 atgctcgttc acggccacac actcgttttg cacaaccaac ttccgttctg gttgaacaga 360
 cagtggacca aagaagaact cctgaaagtc cttgaggacc acatcaaaac agtcgtcgtt 420
 cacttcaaag gaagggtgaa gatttgggac gtggtgaacg aagcggtcag cgacatgggc 480
 agttacagag agaccattttg gtacaagacc atcggaccgg agtacatcga aaaggcattc 540
 gtgtgggcaa gacaagccga tccggaagcg atcctcatat acaacgacta caatatagaa 600
 acgatcaatc ccaaatcgaa tttcacctac cagctcatca aggagctgaa agaaaaaggt 660
 gtgccgatag acggcatcgg ttttcaaatg cacatagaca tcaacggaat aaactatgac 720
 agtttcagaa acaacctgaa gaggttcgct gatctcggtc tgaagctcta catcacggaa 780
 atggtatgtga gaatacccaa gaacgcaact gaaaaagact tggacagaca ggcagaaatc 840
 tacgcaaga tcttcgaaat ctgcttagag aatcctgcgg tccaagccat acagttcttg 900
 ggtttcacgg acaagtattc ctgggtgcct ggctttttca gcgggtacga tcatgcgctg 960
 atctttgaca gggactacag cccaagccc gcgtattttg cgataaagag ggtgctcgaa 1020
 gccaaagtgga gcaagggacg ctga 1044

<210> 150
 <211> 347
 <212> PRT
 <213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(18)

<400> 150

```

Met Lys Lys Leu Phe Val Ala Val Val Leu Leu Pro Leu Ala Thr Phe
1      5      10      15
Phe Ala Ser Asp Gly Leu Glu Gly Glu Pro Leu Arg Ser Leu Ala Glu
20      25      30
Lys Leu Gly Ile Tyr Ile Gly Tyr Ala Ser Ile Asn His Phe Trp Thr
35      40      45
Leu Pro Asp Ser Asn Lys Tyr Thr Glu Val Ala Lys Arg Glu Phe Asn
50      55      60
Ile Leu Thr Pro Glu Asn Gln Met Lys Trp Asp Ser Leu His Pro Glu
65      70      75      80
Pro Asp Arg Tyr Asn Phe Thr Tyr Ala Glu Arg His Val Glu Phe Ala
85      90      95
Leu Glu Asn Asn Met Leu Val His Gly His Thr Leu Val Trp His Asn
100     105     110
Gln Leu Pro Phe Trp Leu Asn Arg Gln Trp Thr Lys Glu Glu Leu Leu
115     120     125
Lys Val Leu Glu Asp His Ile Lys Thr Val Val Gly His Phe Lys Gly
130     135     140
Arg Val Lys Ile Trp Asp Val Val Asn Glu Ala Val Ser Asp Met Gly
145     150     155     160
Ser Tyr Arg Glu Thr Ile Trp Tyr Lys Thr Ile Gly Pro Glu Tyr Ile
165     170     175
Glu Lys Ala Phe Val Trp Ala Arg Gln Ala Asp Pro Glu Ala Ile Leu
180     185     190
Ile Tyr Asn Asp Tyr Asn Ile Glu Thr Ile Asn Pro Lys Ser Asn Phe
195     200     205
Thr Tyr Gln Leu Ile Lys Glu Leu Lys Glu Lys Gly Val Pro Ile Asp
210     215     220
Gly Ile Gly Phe Gln Met His Ile Asp Ile Asn Gly Ile Asn Tyr Asp
225     230     235     240
Ser Phe Arg Asn Asn Leu Lys Arg Phe Ala Asp Leu Gly Leu Lys Leu
245     250     255
Tyr Ile Thr Glu Met Asp Val Arg Ile Pro Lys Asn Ala Thr Glu Lys
260     265     270
Asp Leu Asp Arg Gln Ala Glu Ile Tyr Ala Lys Ile Phe Glu Ile Cys
275     280     285
Leu Glu Asn Pro Ala Val Gln Ala Ile Gln Phe Trp Gly Phe Thr Asp
290     295     300
Lys Tyr Ser Trp Val Pro Gly Phe Phe Ser Gly Tyr Asp His Ala Leu
305     310     315     320
Ile Phe Asp Arg Asp Tyr Ser Pro Lys Pro Ala Tyr Phe Ala Ile Lys
325     330     335
Arg Val Leu Glu Ala Lys Val Ser Lys Gly Arg
340     345

```

<210> 151

<211> 1131

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 151

```

atgcgatcta tgccacttta tgtgttggtta tgcagcgccc ttctgaccgg cagcctatat      60
gcacaagacc aaaatgcttc tttaaaacag gccttttagca aaaacttttag tattggcaca      120
gccttaagtg ctacacaaat tcaggggcaaa gagccgggca cactggaatt ggtaacacag      180
caatttaacg cgggtgacggc agaaaacgtg atgaagtggg aaatcattga acctgtggaa      240
ggccagttca accttgctgc cgccgacgcc atgattgaat tcgccgaagc caatcatatc      300
aaggtgatag gccatgtgct gttatggcac gaacaaacac cagcctgggt atttctggac      360
gccaaaggcc aggccgcctc aaaggaactg gtgttatcac ggctaaaaaa ccatatcaat      420

```

gccgtaatgg	gccgctacaa	aggccgtatt	catggctggg	atgcagtcaa	cgaagcctta	480
aatgaagacg	gcactctgcg	ccaatccaac	tggtataaag	ctttaggcga	cgactatata	540
gccacagtct	ttgaactggc	gcatcaggcc	gacccgaaag	ccgaactcta	ttacaacgac	600
ttcaatttat	ttaaaccgga	aaaacgcgct	ggtgtactca	aactggtggc	agctttaaaa	660
gcgaaaaatg	tgcctatcca	cggcataggc	gagcaaggcc	attacagcct	ggattaccct	720
gagctgcagc	aagtagaaga	ctctattgtg	gcttttaaaa	acactggcct	gaaagtgggtg	780
attaccgaac	tggatatctc	agttttaccc	ttccctgagc	cagaaaagat	tggtgctgat	840
atctcactca	atatgcagtt	aaaacaagaa	cttaatccct	acgccgatgg	cttaccctaaa	900
gaagtcagcg	atcaactgac	agaaaaatac	ctgcaattat	ttcagctatt	tttacgccac	960
agcgacgcca	tcgaacgcgt	gaccttatgg	ggcgtaaacg	acaaccaaac	ctggcgcaac	1020
aactggccaa	tgaaaggcag	aacagactac	cccttactct	tcgaccggaa	aaaccagcca	1080
aaagaagtgg	ttcctgcatt	gattaaactg	gcggaaaaag	ctggtaaata	a	1131

<210> 152

<211> 376

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(21)

<400> 152

Met	Arg	Ser	Met	Pro	Leu	Tyr	Val	Leu	Leu	Cys	Ser	Ala	Leu	Leu	Thr
1				5				10						15	
Gly	Ser	Leu	Tyr	Ala	Gln	Asp	Gln	Asn	Ala	Ser	Leu	Lys	Gln	Ala	Phe
			20					25					30		
Ser	Lys	Asn	Phe	Ser	Ile	Gly	Thr	Ala	Leu	Ser	Ala	Thr	Gln	Ile	Gln
		35					40					45			
Gly	Lys	Glu	Pro	Gly	Thr	Leu	Glu	Leu	Val	Thr	Gln	Gln	Phe	Asn	Ala
	50					55					60				
Val	Thr	Ala	Glu	Asn	Val	Met	Lys	Trp	Glu	Ile	Ile	Glu	Pro	Val	Glu
65					70					75					80
Gly	Gln	Phe	Asn	Phe	Ala	Ala	Ala	Asp	Ala	Met	Ile	Glu	Phe	Ala	Glu
			85					90						95	
Ala	Asn	His	Ile	Lys	Val	Ile	Gly	His	Val	Leu	Leu	Trp	His	Glu	Gln
			100					105					110		
Thr	Pro	Ala	Trp	Val	Phe	Leu	Asp	Ala	Lys	Gly	Gln	Ala	Ala	Ser	Lys
		115					120					125			
Glu	Leu	Val	Leu	Ser	Arg	Leu	Lys	Asn	His	Ile	Asn	Ala	Val	Met	Gly
	130					135					140				
Arg	Tyr	Lys	Gly	Arg	Ile	His	Gly	Trp	Asp	Ala	Val	Asn	Glu	Ala	Leu
145					150					155					160
Asn	Glu	Asp	Gly	Thr	Leu	Arg	Gln	Ser	Asn	Trp	Tyr	Lys	Ala	Leu	Gly
			165						170					175	
Asp	Asp	Tyr	Ile	Ala	Thr	Val	Phe	Glu	Leu	Ala	His	Gln	Ala	Asp	Pro
			180					185					190		
Lys	Ala	Glu	Leu	Tyr	Tyr	Asn	Asp	Phe	Asn	Leu	Phe	Lys	Pro	Glu	Lys
	195						200					205			
Arg	Ala	Gly	Val	Leu	Lys	Leu	Val	Ala	Ala	Leu	Lys	Ala	Lys	Asn	Val
	210					215					220				
Pro	Ile	His	Gly	Ile	Gly	Glu	Gln	Gly	His	Tyr	Ser	Leu	Asp	Tyr	Pro
225					230					235					240
Glu	Leu	Gln	Gln	Val	Glu	Asp	Ser	Ile	Val	Ala	Phe	Lys	Asn	Thr	Gly
			245						250					255	
Leu	Lys	Val	Val	Ile	Thr	Glu	Leu	Asp	Ile	Ser	Val	Leu	Pro	Phe	Pro
			260					265					270		
Glu	Pro	Glu	Lys	Ile	Gly	Ala	Asp	Ile	Ser	Leu	Asn	Met	Gln	Leu	Lys
		275					280					285			
Gln	Glu	Leu	Asn	Pro	Tyr	Ala	Asp	Gly	Leu	Pro	Lys	Glu	Val	Ser	Asp
	290					295					300				
Gln	Leu	Thr	Glu	Lys	Tyr	Leu	Gln	Leu	Phe	Gln	Leu	Phe	Leu	Arg	His
305					310					315					320
Ser	Asp	Ala	Ile	Glu	Arg	Val	Thr	Leu	Trp	Gly	Val	Asn	Asp	Asn	Gln
			325						330					335	
Thr	Trp	Arg	Asn	Asn	Trp	Pro	Met	Lys	Gly	Arg	Thr	Asp	Tyr	Pro	Leu
			340					345					350		

Leu Phe Asp Arg Lys Asn Gln Pro Lys Glu Val Val Pro Ala Leu Ile
 355 360
 Lys Leu Ala Glu Lys Ala Gly Lys
 370 375

<210> 153
 <211> 1020
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 153
 atgggtgcta tgggcctggc ggcgctgtat tcgctgccag ccaatgcaca gacctgcatt 60
 acgcagagtc agacgggcac caacaacggc cactattttt cgttctggaa ggacaatccg 120
 ggaacgggtca atttctgtat gtatgccaac ggccgttaca cgtctaactg gaacggcatc 180
 aacaattggg tcggcggcaa aggttggaac accggctcgc gcagaaacgt cacctactct 240
 ggctcgttca actctcccgg caatggctat ctggctgctc tactggctgg accaccaatc 300
 ctgttggtcg agtactacat catcgagagc tggggaaatt ggcgcccgcc gggttcggat 360
 ggaacattgt taggcaccgt cactagcgac ggcggtaact acgatatcta tcgctcgcgc 420
 cgcaccaacg cgccttgtat cactggcaac tcctgtaact tcgatcagta ctggagcgta 480
 cggcaatcca agcgcgtggg cggcacgatt accacgggca atcacttcga cgcttgggcg 540
 gcacgcggct tgaacctcgg caccgacaac taccaagtga tggcgaccga gggatatcag 600
 agcaacggca gctccgacat caccattagc gacaacccgg gaccgacgcc aggaccact 660
 ccgaaccgga atcccacgcc gggcaccaag aatttcacgg tgcgcgcgcg cggaaaccgcg 720
 ggggggtgagt ccatcacgct gcgtgtgaac aatcagaacg tgcagacctg gacgctgtcg 780
 accagctacc agaacttcac ggcgtccacg acgttgagtg gtggcatcac ggtcgcgttc 840
 accaatgatg gtggtagtgc agacgttcag gtggattaca tccaggtgaa cggcgcaact 900
 cgacaatccg agagccagac gtacaacacc ggcctctatg ccaacggcag ttgcggcggc 960
 ggctcgaaca gcgagtggat gcattgcaat ggagcgatcg gctacggcaa cacgccgtag 1020

<210> 154
 <211> 339
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(16)

<400> 154
 Met Gly Ala Met Gly Leu Ala Ala Leu Tyr Ser Leu Pro Ala Asn Ala
 1 5 10 15
 Gln Thr Cys Ile Thr Gln Ser Gln Thr Gly Thr Asn Asn Gly His Tyr
 20 25 30
 Phe Ser Phe Trp Lys Asp Asn Pro Gly Thr Val Asn Phe Cys Met Tyr
 35 40 45
 Ala Asn Gly Arg Tyr Thr Ser Asn Trp Asn Gly Ile Asn Asn Trp Val
 50 55 60
 Gly Gly Lys Gly Trp Gln Thr Gly Ser Arg Arg Asn Val Thr Tyr Ser
 65 70 75 80
 Gly Ser Phe Asn Ser Pro Gly Asn Gly Tyr Leu Ala Ala Leu Leu Ala
 85 90 95
 Gly Pro Pro Ile Leu Leu Val Glu Tyr Tyr Ile Ile Glu Ser Trp Gly
 100 105 110
 Asn Trp Arg Pro Pro Gly Ser Asp Gly Thr Leu Leu Gly Thr Val Thr
 115 120 125
 Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Ser Arg Arg Thr Asn Ala
 130 135 140
 Pro Cys Ile Thr Gly Asn Ser Cys Asn Phe Asp Gln Tyr Trp Ser Val
 145 150 155 160
 Arg Gln Ser Lys Arg Val Gly Gly Thr Ile Thr Thr Gly Asn His Phe
 165 170 175
 Asp Ala Trp Ala Ala Arg Gly Leu Asn Leu Gly Thr His Asn Tyr Gln
 180 185 190
 Val Met Ala Thr Glu Gly Tyr Gln Ser Asn Gly Ser Ser Asp Ile Thr

Ile	Ser	195	Asp	Asn	Pro	Gly	Pro	200	Thr	Pro	Gly	Pro	Thr	205	Pro	Asn	Pro	Asn
210	210						215						220					
Pro	Thr	Pro	Gly	Thr	Lys	Asn	Phe	Thr	Val	Arg	Ala	Arg	Gly	Thr	Ala			
225					230					235					240			
Gly	Gly	Glu	Ser	Ile	Thr	Leu	Arg	Val	Asn	Asn	Gln	Asn	Val	Gln	Thr			
				245					250					255				
Trp	Thr	Leu	Ser	Thr	Ser	Tyr	Gln	Asn	Phe	Thr	Ala	Ser	Thr	Thr	Leu			
		260						265						270				
Ser	Gly	Gly	Ile	Thr	Val	Ala	Phe	Thr	Asn	Asp	Gly	Gly	Ser	Arg	Asp			
		275						280						285				
Val	Gln	Val	Asp	Tyr	Ile	Gln	Val	Asn	Gly	Ala	Thr	Arg	Gln	Ser	Glu			
		290				295					300							
Ser	Gln	Thr	Tyr	Asn	Thr	Gly	Leu	Tyr	Ala	Asn	Gly	Ser	Cys	Gly	Gly			
305					310					315					320			
Gly	Ser	Asn	Ser	Glu	Trp	Met	His	Cys	Asn	Gly	Ala	Ile	Gly	Tyr	Gly			
			325						330					335				

Asn Thr Pro

<210> 155
 <211> 1836
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 155																		
atgaaaggat	taattgcggc	agcgccttgct	ggcttggcat	tcggggcctc	cctatcctgg													60
ggacagtgc	caacgtttac	caccagtacc	attcagaatt	gtaatggcat	tgattacgag													120
ctctggagtc	agaataacaa	gggcaccgta	agcatgaaga	ttacgggagg	gagcacgaat													180
ccgaatggag	gaactttcga	tgctacctgg	aatggcaccg	agaatatcct	ggctagagct													240
ggtaagaaat	ggggctcgtc	cagcactacc	accccacgt	ccgcaggcaa	tattactctt													300
gaattcgcgg	cgacatggtc	ctcaagcgat	aacgtaaaaa	tgcttggagt	ctatggctgg													360
gcgtactatc	caactggaag	tatcccgact	aaacaggaaa	atggagcaag	tacctcattc													420
acaaatcaaa	ttgagtacta	catcatccag	gatcgtggta	gctataatgc	tgcatcgggt													480
ggaacgaact	ccaaaaaata	cggcgaaggg	acgatcgatg	gaattctgta	tgaattctat													540
atcgcacaca	gaatcaacca	gcctgatctg	tcaggaaaaga	gtggaaactt	caagcaatac													600
ttcagcgtcc	cgaaaagtac	gagcagccat	aggcaaagtg	ggacgattac	cgtttccaaa													660
catttccagg	cctgggaaaa	tgccggaatg	aaaatgatgt	cctgtcgctt	gtatgaagtc													720
gcaatgaaag	tcgagtccta	taccggttct	gcgaccgggt	ttggctctgc	gaaggttaca													780
aagaatatac	tcaccattgg	tggaatcttg	agcagtagca	gtactgcaag	cagtagcagc													840
acagtaagta	cagtagcag	caatgcatat	acgcttgcca	cgaatgtttc	tcccgtgga													900
gccggaacag	tgaccaggag	cccccaatac	gcgacctatg	ccccgaatgc	ttcagtagac													960
cttactgcaa	cgccgagtag	cggttggaaa	tttgtcggtt	gggctgggga	tcttacgtca													1020
actacgagta	ctgctaccgt	caccatgacc	aaagatatta	ccgcaactgc	aaaatttgaa													1080
ctggtatcgg	gagatggcac	gaccaacttg	atcaaggatg	gaaacttccc	cagtagcagc													1140
gtcatctcca	caggtgatgg	cacctcctgg	aaagctcggt	aaggtagaaa	ctggggtaat													1200
tccgcagcaa	cgacgagtgt	cagcaatgga	atcgcgactg	tcaatgtgac	caccattgga													1260
tctcaaacct	atcaacccca	gctaattcag	tataacgtgg	ctctttacaa	ggatattgagc													1320
tacaagctca	ccttcaaggg	aaaagctgct	gctgcaagga	aaattgaagt	cgatttccaa													1380
cagtcggtgg	acccatgggc	tgatattgct	tccaagggaat	tcgatcttac	aacgacagag													1440
cagacatatg	agttcgtatt	taaaatgact	agcgtactg	acacggcttc	acagttcgcg													1500
ttcaatctcg	gccaggcaac	aggcgccgtc	aatattagtg	atgtaaagct	agtatatacg													1560
acagctggta	caacacccgt	attccgtgga	tataatgagg	cggaacaca	ggagaggcct													1620
gtattcatat	ccttggatgg	taggacgttg	aacattgttc	cagtgtatgg	agccaaactg													1680
caggtcaagt	tagtgacatg	caatggtaag	atgagagcct	ccttcaatgt	ggtcggaatt													1740
gcttccatcc	cgctgtccaa	tatccccgct	gggcgggtatt	atattgacgt	aagtggtagc													1800
ggcgtaagc	aggcatcccc	gatagttctg	gaataa															1836

<210> 156
 <211> 611
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(21)

<400> 156

Met	Lys	Gly	Leu	Ile	Ala	Ala	Ala	Leu	Ala	Gly	Leu	Ala	Phe	Gly	Ala
1				5					10					15	
Ser	Leu	Ser	Trp	Gly	Gln	Cys	Thr	Thr	Phe	Thr	Thr	Ser	Thr	Ile	Gln
			20					25					30		
Asn	Cys	Asn	Gly	Ile	Asp	Tyr	Glu	Leu	Trp	Ser	Gln	Asn	Asn	Lys	Gly
		35					40					45			
Thr	Val	Ser	Met	Lys	Ile	Thr	Gly	Gly	Ser	Thr	Asn	Pro	Asn	Gly	Gly
	50					55					60				
Thr	Phe	Asp	Ala	Thr	Trp	Asn	Gly	Thr	Glu	Asn	Ile	Leu	Ala	Arg	Ala
65					70					75				80	
Gly	Lys	Lys	Trp	Gly	Ser	Ser	Ser	Thr	Thr	Thr	Pro	Thr	Ser	Ala	Gly
			85						90					95	
Asn	Ile	Thr	Leu	Glu	Phe	Ala	Ala	Thr	Trp	Ser	Ser	Ser	Asp	Asn	Val
			100					105					110		
Lys	Met	Leu	Gly	Val	Tyr	Gly	Trp	Ala	Tyr	Tyr	Pro	Thr	Gly	Ser	Ile
	115						120				125				
Pro	Thr	Lys	Gln	Glu	Asn	Gly	Ala	Ser	Thr	Ser	Phe	Thr	Asn	Gln	Ile
	130					135					140				
Glu	Tyr	Tyr	Ile	Ile	Gln	Asp	Arg	Gly	Ser	Tyr	Asn	Ala	Ala	Ser	Gly
145					150					155				160	
Gly	Thr	Asn	Ser	Lys	Lys	Tyr	Gly	Glu	Gly	Thr	Ile	Asp	Gly	Ile	Leu
				165					170					175	
Tyr	Glu	Phe	Tyr	Ile	Ala	Asp	Arg	Ile	Asn	Gln	Pro	Asp	Leu	Ser	Gly
			180					185					190		
Lys	Ser	Gly	Asn	Phe	Lys	Gln	Tyr	Phe	Ser	Val	Pro	Lys	Ser	Thr	Ser
		195					200					205			
Ser	His	Arg	Gln	Ser	Gly	Thr	Ile	Thr	Val	Ser	Lys	His	Phe	Gln	Ala
	210					215					220				
Trp	Glu	Asn	Ala	Gly	Met	Lys	Met	Met	Ser	Cys	Arg	Leu	Tyr	Glu	Val
225					230					235				240	
Ala	Met	Lys	Val	Glu	Ser	Tyr	Thr	Gly	Ser	Ala	Thr	Gly	Val	Gly	Ser
			245						250					255	
Ala	Lys	Val	Thr	Lys	Asn	Ile	Leu	Thr	Ile	Gly	Gly	Ile	Leu	Ser	Ser
			260					265					270		
Ser	Ser	Thr	Ala	Ser	Ser	Ser	Ser	Thr	Val	Ser	Ser	Ser	Ser	Ser	Asn
		275					280					285			
Ala	Tyr	Thr	Leu	Val	Thr	Asn	Val	Ser	Pro	Ala	Gly	Ala	Gly	Thr	Val
	290					295					300				
Thr	Arg	Ser	Pro	Asn	Thr	Ala	Thr	Tyr	Ala	Pro	Asn	Ala	Ser	Val	Gln
305					310					315					320
Leu	Thr	Ala	Thr	Pro	Ser	Thr	Gly	Trp	Lys	Phe	Val	Gly	Trp	Ala	Gly
				325					330					335	
Asp	Leu	Thr	Ser	Thr	Thr	Ser	Thr	Ala	Thr	Val	Thr	Met	Thr	Lys	Asp
			340					345					350		
Ile	Thr	Ala	Thr	Ala	Lys	Phe	Glu	Leu	Val	Ser	Gly	Asp	Gly	Thr	Thr
		355					360					365			
Asn	Leu	Ile	Lys	Asp	Gly	Asn	Phe	Pro	Ser	Ser	Ser	Val	Ile	Ser	Thr
	370					375					380				
Gly	Asp	Gly	Thr	Ser	Trp	Lys	Leu	Gly	Gln	Gly	Thr	Asn	Trp	Gly	Asn
385					390					395					400
Ser	Ala	Ala	Thr	Thr	Ser	Val	Ser	Asn	Gly	Ile	Ala	Thr	Val	Asn	Val
				405					410					415	
Thr	Thr	Ile	Gly	Ser	Gln	Thr	Tyr	Gln	Pro	Gln	Leu	Ile	Gln	Tyr	Asn
			420					425					430		
Val	Ala	Leu	Tyr	Lys	Asp	Met	Ser	Tyr	Lys	Leu	Thr	Phe	Lys	Ala	Lys
		435					440					445			
Ala	Ala	Ala	Ala	Arg	Lys	Ile	Glu	Val	Ala	Phe	Gln	Gln	Ser	Val	Asp
	450					455					460				
Pro	Trp	Ala	Gly	Tyr	Ala	Ser	Lys	Glu	Phe	Asp	Leu	Thr	Thr	Thr	Glu
465					470					475					480
Gln	Thr	Tyr	Glu	Phe	Val	Phe	Lys	Met	Thr	Ser	Ala	Thr	Asp	Thr	Ala
				485					490					495	
Ser	Gln	Phe	Ala	Phe	Asn	Leu	Gly	Gln	Ala	Thr	Gly	Ala	Val	Asn	Ile
			500					505					510		
Ser	Asp	Val	Lys	Leu	Val	Tyr	Thr	Thr	Ala	Gly	Thr	Thr	Pro	Val	Phe


```

      515      520      525
Arg Gly Tyr Asn Glu Ala Ala Thr Gln Glu Arg Pro Val Phe Ile Ser
   530   535   540
Leu Asp Gly Arg Thr Leu Asn Ile Val Pro Val Tyr Gly Ala Lys Leu
545   550   555   560
Gln Val Lys Leu Val Asp Ile Asn Gly Lys Met Arg Ala Ser Phe Asn
      565      570      575
Val Val Gly Ile Ala Ser Ile Pro Leu Ser Asn Ile Pro Ala Gly Arg
      580      585      590
Tyr Tyr Ile Asp Val Ser Gly Asp Gly Val Lys Gln Ala Ser Pro Ile
   595   600   605
Val Leu Glu
   610

```

<210> 157
 <211> 645
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

```

<400> 157
atgtttaagt taagtaagaa aattttgatg gtgttattaa caatttcaat gagttttatt      60
agcttatttg cagtaaccgc gtatgcagct tgcacagact actggcaaaa ttggactgat      120
ggtggtggga cagtaaatgc taccaatgga tctgatggca attacagtgt ttcatgggtca      180
aattgcggga attttgttgt tggtaaaggc tggactaccg gatcagcaac tagggtaata      240
aactataatg ccggagcctt ttcgccgtcc ggcaatggat atttagctct ttatgggtgg      300
acgagaaatt cactcataga atattacgtc gttgatagct gggggactta tagacctact      360
ggaacttata aaggcactgt gactagtgat ggagggacat atgacatata cacgactaca      420
cgaaccaacg caccttccat tgacggcaat aatacaaatt tcacccagtt ctggagtgtt      480
aggcagtaaa agagaccgat tggtaaccaac aataccatca cttttagcaa ccacgttaac      540
gcctggaaga gtaaaggaat gaatctgggg agtagttggg cttatcaggt attagcgaca      600
gagggatatc aaagtagtgg gtactctaac gtaacggtct ggtaa      645

```

<210> 158
 <211> 214
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(29)

```

<400> 158
Met Phe Lys Leu Ser Lys Lys Ile Leu Met Val Leu Leu Thr Ile Ser
 1      5      10      15
Met Ser Phe Ile Ser Leu Phe Ala Val Thr Ala Tyr Ala Ala Ser Thr
      20      25      30
Asp Tyr Trp Gln Asn Trp Thr Asp Gly Gly Gly Thr Val Asn Ala Thr
      35      40      45
Asn Gly Ser Asp Gly Asn Tyr Ser Val Ser Trp Ser Asn Cys Gly Asn
      50      55      60
Phe Val Val Gly Lys Gly Trp Thr Thr Gly Ser Ala Thr Arg Val Ile
      65      70      75      80
Asn Tyr Asn Ala Gly Ala Phe Ser Pro Ser Gly Asn Gly Tyr Leu Ala
      85      90      95
Leu Tyr Gly Trp Thr Arg Asn Ser Leu Ile Glu Tyr Tyr Val Val Asp
      100      105      110
Ser Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys Gly Thr Val Thr
      115      120      125
Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr Thr Arg Thr Asn Ala
      130      135      140
Pro Ser Ile Asp Gly Asn Asn Thr Asn Phe Thr Gln Phe Trp Ser Val
      145      150      155      160
Arg Gln Ser Lys Arg Pro Ile Gly Thr Asn Asn Thr Ile Thr Phe Ser
      165      170      175

```

Asn His Val Asn Ala Trp Lys Ser Lys Gly Met Asn Leu Gly Ser Ser
 180 185 190
 Trp Ala Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln Ser Ser Gly Tyr
 195 200 205
 Ser Asn Val Thr Val Trp
 210

<210> 159
 <211> 1041
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 159
 atgatcagtc tcaaacgagt ggcggcgctc ctgtgcgtcg caggctctggg catgtctgcg 60
 gcaaacgcgc agacctgcct cacgtcgagt caaacggca ctaacaatgg cttctattat 120
 tccttctgga aggacagtcc gggcacggtg aatttttgcc tgcagtcagg cgcccggtac 180
 acatcgaact ggagcggcat caacaactgg gtgggcggca agggatggca gaccgggtca 240
 cgccgggaaca tcacgtactc gggcagcttc aattcacggg gcaacggcta cctggcgctt 300
 tacggatgga ccaccaatcc actcgtcgag tactacgtcg tcgatatgctg ggggagctgg 360
 cgtccgcagg gttcggacgg aacgttctcg gggacggtca acagcgatgg cggaacgtat 420
 gacatctatc gcgcgcagcg ggtcaacgag ccgtccatca tcggcaacgc cacgttctat 480
 caatactgga gcgttcggca gtcgaagcgg gtaggtggga cgatcaccac cggaaccac 540
 ttccagcgcgt ggccacgctt gggcctgaac ctgggcaact acaactacca gatcatggcg 600
 accgagggct accaaagcag cggcagctcc gacatcacgg tgagtgaagg cggtagcagc 660
 agtgggtggcg gaagcagcac gagcagcagc agcggcggtg gtggcaccaa gagcttcacg 720
 gttcgtgcgc gcggtaccgc gggcggtgag tccatcacgc tgcgcgtgaa caaccagaac 780
 gtgcagacct ggacgctggg caccagcatg acgaactaca cggcgctcgac ttactgagc 840
 ggcggcatca ccgtggtgta cacgaacgac agcggtaacc gcgacgtgca ggtggactac 900
 atcgtctgta acggccagac gcgccagtcc gaagcccaga gctacaacac cggcctttat 960
 gcgaacgggc gttgcggcgg tggctccaac agcgaatgga tgcattgcaa cggcgccatc 1020
 ggctacggca atacaccgta a 1041

<210> 160
 <211> 346
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(23)

<400> 160
 Met Ile Ser Leu Lys Arg Val Ala Ala Leu Leu Cys Val Ala Gly Leu
 1 5 10 15
 Gly Met Ser Ala Ala Asn Ala Gln Thr Cys Leu Thr Ser Ser Gln Thr
 20 25 30
 Gly Thr Asn Asn Gly Phe Tyr Tyr Ser Phe Trp Lys Asp Ser Pro Gly
 35 40 45
 Thr Val Asn Phe Cys Leu Gln Ser Gly Gly Arg Tyr Thr Ser Asn Trp
 50 55 60
 Ser Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Gln Thr Gly Ser
 65 70 75 80
 Arg Arg Asn Ile Thr Tyr Ser Gly Ser Phe Asn Ser Pro Gly Asn Gly
 85 90 95
 Tyr Leu Ala Leu Tyr Gly Trp Thr Thr Asn Pro Leu Val Glu Tyr Tyr
 100 105 110
 Val Val Asp Ser Trp Gly Ser Trp Arg Pro Pro Gly Ser Asp Gly Thr
 115 120 125
 Phe Leu Gly Thr Val Asn Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Arg
 130 135 140
 Ala Gln Arg Val Asn Ala Pro Ser Ile Ile Gly Asn Ala Thr Phe Tyr
 145 150 155 160
 Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg Val Gly Gly Thr Ile Thr
 165 170 175

Thr Gly Asn His Phe Asp Ala Trp Ala Ser Val Gly Leu Asn Leu Gly
 180 185 190
 Thr His Asn Tyr Gln Ile Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly
 195 200 205
 Ser Ser Asp Ile Thr Val Ser Glu Gly Gly Ser Ser Ser Gly Gly Gly
 210 215 220
 Ser Ser Thr Ser Ser Ser Ser Gly Gly Gly Gly Thr Lys Ser Phe Thr
 225 230 235 240
 Val Arg Ala Arg Gly Thr Ala Gly Gly Glu Ser Ile Thr Leu Arg Val
 245 250 255
 Asn Asn Gln Asn Val Gln Thr Trp Thr Leu Gly Thr Ser Met Thr Asn
 260 265 270
 Tyr Thr Ala Ser Thr Ser Leu Ser Gly Gly Ile Thr Val Val Tyr Thr
 275 280 285
 Asn Asp Ser Gly Asn Arg Asp Val Gln Val Asp Tyr Ile Val Val Asn
 290 295 300
 Gly Gln Thr Arg Gln Ser Glu Ala Gln Ser Tyr Asn Thr Gly Leu Tyr
 305 310 315 320
 Ala Asn Gly Arg Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His Cys
 325 330 335
 Asn Gly Ala Ile Gly Tyr Gly Asn Thr Pro
 340 345

<210> 161
 <211> 1047
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 161
 atgttcaaag gtcttttgaa atcgggtcctc accggcaagc gagccggtgc ggtgttcatc 60
 tgtctggccg gactgtggat gacacaggcg caggcgacaga cgtgcatcgg ttcaccacaa 120
 acgggcaaca acggcggtctt cttcttttcg ttctggaaag acaatccggg gtcggtgaat 180
 ttctgcatgt actccggcgg tcgctatacc tccagctgga gcggcatcaa caactgggta 240
 ggtgggaagg gctggcaaac cggttcatcc cgcacggtga cgtattcggg cacgttcaac 300
 tcgcccggaa acggctacct gactctgtac ggatggacca ccaatccgct ggtcgagtac 360
 tacatcgtgg acagctgggg cagctaccgt ccgcttgag gccagggctt catgggcacg 420
 gtcaccagcg acggcggaac gtatgacatc taccgggttc gccgcaccaa tgcgccgtgc 480
 atcacaggca acaactgcaa cttcgaccag tactggagcg tgcgtcagtc gaggcgggtg 540
 ggccggcacca tcaccaccgc caaccatttc aacgcgtggc gtacgctcgg catgaatctc 600
 gggcagcaca actaccaggt gatggcgacc gaaggattcc agagcagtgg cagctcggac 660
 atcacctgta gcgaaggatc tggcggtggc ggcggaggtg gcggcggtgg caccaagagc 720
 ttcacggtgc gcgcgcgcgg caccgcgggc ggcgagtcca tcacgctgcg cgtcaacaac 780
 caggtcgtgc agagctggac cttgagcacc agcatgcaga actacacggc ctcgaccacg 840
 atgagcggcg gcatcacggt gaacttcacc aacgacggca ccaaccgcga cgtgcagggtg 900
 gactacatca tcgtgaatgg ccagacgcgt cagtccgaag cgcagacgta caacaccggg 960
 ctgtacgcca acggccgttg cgggtggcgg tcgaacagcg agtggatgca ttgcaatggc 1020
 gcgatcgggt acggcgacac gccctga 1047

<210> 162
 <211> 348
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(32)

<400> 162
 Met Phe Lys Gly Leu Leu Lys Ser Val Leu Thr Gly Lys Arg Ala Gly
 1 5 10 15
 Ala Val Phe Ile Cys Leu Ala Gly Leu Trp Met Thr Gln Ala Gln Ala
 20 25 30
 Gln Thr Cys Ile Gly Ser Pro Gln Thr Gly Asn Asn Gly Gly Phe Phe
 35 40 45

Phe Ser Phe Trp Lys Asp Asn Pro Gly Ser Val Asn Phe Cys Met Tyr
 50 55 60
 Ser Gly Gly Arg Tyr Thr Ser Ser Trp Ser Gly Ile Asn Asn Trp Val
 65 70 75 80
 Gly Gly Lys Gly Trp Gln Thr Gly Ser Ser Arg Thr Val Thr Tyr Ser
 85 90 95
 Gly Thr Phe Asn Ser Pro Gly Asn Gly Tyr Leu Thr Leu Tyr Gly Trp
 100 105 110
 Thr Thr Asn Pro Leu Val Glu Tyr Tyr Ile Val Asp Ser Trp Gly Ser
 115 120 125
 Tyr Arg Pro Pro Gly Gly Gln Gly Phe Met Gly Thr Val Thr Ser Asp
 130 135 140
 Gly Gly Thr Tyr Asp Ile Tyr Arg Val Arg Arg Thr Asn Ala Pro Cys
 145 150 155 160
 Ile Thr Gly Asn Asn Cys Asn Phe Asp Gln Tyr Trp Ser Val Arg Gln
 165 170 175
 Ser Arg Arg Val Gly Gly Thr Ile Thr Thr Ala Asn His Phe Asn Ala
 180 185 190
 Trp Arg Thr Leu Gly Met Asn Leu Gly Gln His Asn Tyr Gln Val Met
 195 200 205
 Ala Thr Glu Gly Phe Gln Ser Ser Gly Ser Ser Asp Ile Thr Val Ser
 210 215 220
 Glu Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Thr Lys Ser
 225 230 235 240
 Phe Thr Val Arg Ala Arg Gly Thr Ala Gly Gly Glu Ser Ile Thr Leu
 245 250 255
 Arg Val Asn Asn Gln Val Val Gln Ser Trp Thr Leu Ser Thr Ser Met
 260 265 270
 Gln Asn Tyr Thr Ala Ser Thr Thr Met Ser Gly Gly Ile Thr Val Asn
 275 280 285
 Phe Thr Asn Asp Gly Thr Asn Arg Asp Val Gln Val Asp Tyr Ile Ile
 290 295 300
 Val Asn Gly Gln Thr Arg Gln Ser Glu Ala Gln Thr Tyr Asn Thr Gly
 305 310 315 320
 Leu Tyr Ala Asn Gly Arg Cys Gly Gly Gly Ser Asn Ser Glu Trp Met
 325 330 335
 His Cys Asn Gly Ala Ile Gly Tyr Gly Asp Thr Pro
 340 345

<210> 163
 <211> 1068
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 163
 atgaaagcaa agagaatgaa gttgtttgcc gcattttttac tctgtttttac gcttgcaactt 60
 cctggggcag tgcattgcgca gacgatcacc agcaattcgg tcggtacgca tgacggttat 120
 gactatgaat actggaagga cagcgggaat ggaactatgg ttctcggtag tggcggtag 180
 ttcagtgccg agtgagcaaa tatcaataat attctgttcc gtaaaggcaa gaagtcaat 240
 gagacgcaga cccatcagca aattggaaac atttcataaa cctatgggtgc cacctaccaa 300
 ccgaatggca attcgtatatt aacgggtctat ggctggacgg ttgacccctt cgtcgaatat 360
 tacattgtcg atagctgggg cagctggcgt ccgcctggag catcgccaaa ggggactgtt 420
 aacgttgacg gaggaacgta tgacatttat gagacaactc gtgtcaacca gccttccatt 480
 aaaggcacgg caaccttcaa gcagtattgg agtgtccgga cgtcaaaaac gacgagcggg 540
 accatatctg taagcgagca ctttaaggcc tgggagaaat tggggatgac catgggcaag 600
 atgtatgaag tcgcgcttac ggttgaaggc tatcaaagca gtggaagcgc taatgtgtat 660
 agccatacac tgacgatcgg cgggggaaca acacctccac caaccacagg cacaaagatc 720
 gaagccgaga gtatgaccaa aagcggacaa tacactggga atatcagctc gccgttcaac 780
 ggagtcgctt tgtatgccaa caatgattcc gtgaaattca cgcataattt cagcaccggc 840
 acccataact tctcactccg gggggcatca aacaactcca atatggcccc ggttgacctg 900
 aaaatcggcg ggcagacgaa ggggaccttc tatttcggcg gaagcagccc tgcggtctat 960
 actctgaata atgtcagcca tgggaaccgga aatcaagagg ttgaactcgt tgtaaccgac 1020
 gataacggaa catgggatgc ttctattgat tatctcgaga tccattaa 1068

<210> 164
 <211> 355

<212> PRT
<213> Unknown

<220>
<223> obtained from an environmental sample

<221> SIGNAL
<222> (1)...(26)

<400> 164
Met Lys Ala Lys Arg Met Lys Leu Phe Ala Ala Phe Leu Leu Cys Phe
1 5 10 15
Thr Leu Ala Leu Pro Gly Ala Val His Ala Gln Thr Ile Thr Ser Asn
20 25 30
Ser Val Gly Thr His Asp Gly Tyr Asp Tyr Glu Tyr Trp Lys Asp Ser
35 40 45
Gly Asn Gly Thr Met Val Leu Gly Ser Gly Gly Thr Phe Ser Ala Glu
50 55 60
Trp Ser Asn Ile Asn Asn Ile Leu Phe Arg Lys Gly Lys Lys Phe Asn
65 70 75 80
Glu Thr Gln Thr His Gln Gln Ile Gly Asn Ile Ser Ile Thr Tyr Gly
85 90 95
Ala Thr Tyr Gln Pro Asn Gly Asn Ser Tyr Leu Thr Val Tyr Gly Trp
100 105 110
Thr Val Asp Pro Leu Val Glu Tyr Tyr Ile Val Asp Ser Trp Gly Ser
115 120 125
Trp Arg Pro Pro Gly Ala Ser Pro Lys Gly Thr Val Asn Val Asp Gly
130 135 140
Gly Thr Tyr Asp Ile Tyr Glu Thr Thr Arg Val Asn Gln Pro Ser Ile
145 150 155 160
Lys Gly Thr Ala Thr Phe Lys Gln Tyr Trp Ser Val Arg Thr Ser Lys
165 170 175
Arg Thr Ser Gly Thr Ile Ser Val Ser Glu His Phe Lys Ala Trp Glu
180 185 190
Lys Leu Gly Met Thr Met Gly Lys Met Tyr Glu Val Ala Leu Thr Val
195 200 205
Glu Gly Tyr Gln Ser Ser Gly Ser Ala Asn Val Tyr Ser His Thr Leu
210 215 220
Thr Ile Gly Gly Gly Thr Thr Pro Pro Pro Thr Thr Gly Thr Lys Ile
225 230 235 240
Glu Ala Glu Ser Met Thr Lys Ser Gly Gln Tyr Thr Gly Asn Ile Ser
245 250 255
Ser Pro Phe Asn Gly Val Ala Leu Tyr Ala Asn Asn Asp Ser Val Lys
260 265 270
Phe Thr His Asn Phe Thr Thr Gly Thr His Asn Phe Ser Leu Arg Gly
275 280 285
Ala Ser Asn Asn Ser Asn Met Ala Arg Val Asp Leu Lys Ile Gly Gly
290 295 300
Gln Thr Lys Gly Thr Phe Tyr Phe Gly Gly Ser Ser Pro Ala Val Tyr
305 310 315 320
Thr Leu Asn Asn Val Ser His Gly Thr Gly Asn Gln Glu Val Glu Leu
325 330 335
Val Val Thr Ala Asp Asn Gly Thr Trp Asp Ala Phe Ile Asp Tyr Leu
340 345 350
Glu Ile His
355

<210> 165
<211> 1047
<212> DNA
<213> Unknown

<220>
<223> obtained from an environmental sample

<400> 165
gtggggcgca ggagcgccgc cacggcattc atcggcctgg cagcgctgtg tgcctcggcc 60
gccaacgcgc agacctgtct gagctcgagt cagaccggca ccaacaacgg cttctactat 120
tcgttctgga ccgacggcgg tggctccgtg cagttctgcc tgcaatccgc cgggcgctac 180

acctccagct	ggagcaatgt	cggaaactgg	gtcgggtggca	agggctggca	gaccggcgcg	240
cgccgcaaca	tcaactattc	cggcagcttc	aatccctcgg	gtaacgcgta	cctggccgctc	300
tatggctgga	ccacgaatcc	cctgggtggag	tactacatcg	tcgacaactg	gggtacctat	360
cgtccaccgg	gtgggcaggg	attcatgggc	acggttggtca	gcgatggcgg	cacctacgac	420
gtctaccgca	cgcaacgggt	caacgcgccc	tccattcagg	gcaacgcgac	cttctaccag	480
tactggagcg	ttcgccagtc	gaagcgcacc	ggtggaacca	tctccaccgg	caaccatttc	540
gacggctggg	cgacgttcgg	catgaacctg	ggaaccttca	attaccagat	cgtggcgacc	600
gagggctacc	agagcagcgg	caattccgac	atcacggtga	gcgatggcgg	cagcagctcc	660
tcgtcctcca	gcagcagcag	ttcgtcgtcc	tccagcagcg	gcggtggcgg	caccaagagc	720
ttcacggtgc	gcgcgcgcgg	cacggccgga	ggcgagtcga	tcagcctgcg	ggtcaacaac	780
accaacgtgc	agacctggtc	gctgaccacc	agctaccaga	atctcacggc	ctcgaccacg	840
ctgaccggcg	gcatcacctg	caactacacc	aacgacagca	gcggtcacga	cgtacagggtg	900
gactacatca	tcgtgaacgg	ccagaccgcg	cagtccgagg	cgcagagcta	caacaccgga	960
ctctatgcca	acggggcgctg	cgggtgggtg	ggctacagcg	agtggatgca	ttgcaacggc	1020
gccatcggct	acggcaatac	gccgtaa				1047

<210> 166

<211> 348

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(23)

<400> 166

Val	Gly	Arg	Arg	Ser	Ala	Ala	Thr	Ala	Phe	Ile	Gly	Leu	Ala	Ala	Leu
1				5				10						15	
Cys	Ala	Ser	Ala	Ala	Asn	Ala	Gln	Thr	Cys	Leu	Ser	Ser	Ser	Gln	Thr
			20				25					30			
Gly	Thr	Asn	Asn	Gly	Phe	Tyr	Tyr	Ser	Phe	Trp	Thr	Asp	Gly	Gly	Gly
		35					40					45			
Ser	Val	Gln	Phe	Cys	Leu	Gln	Ser	Ala	Gly	Arg	Tyr	Thr	Ser	Ser	Trp
	50					55					60				
Ser	Asn	Val	Gly	Asn	Trp	Val	Gly	Gly	Lys	Gly	Trp	Gln	Thr	Gly	Ala
65					70				75						80
Arg	Arg	Asn	Ile	Asn	Tyr	Ser	Gly	Ser	Phe	Asn	Pro	Ser	Gly	Asn	Ala
			85						90					95	
Tyr	Leu	Ala	Val	Tyr	Gly	Trp	Thr	Thr	Asn	Pro	Leu	Val	Glu	Tyr	Tyr
		100						105					110		
Ile	Val	Asp	Asn	Trp	Gly	Thr	Tyr	Arg	Pro	Pro	Gly	Gly	Gln	Gly	Phe
		115					120					125			
Met	Gly	Thr	Val	Val	Ser	Asp	Gly	Gly	Thr	Tyr	Asp	Val	Tyr	Arg	Thr
	130					135					140				
Gln	Arg	Val	Asn	Ala	Pro	Ser	Ile	Gln	Gly	Asn	Ala	Thr	Phe	Tyr	Gln
145					150					155					160
Tyr	Trp	Ser	Val	Arg	Gln	Ser	Lys	Arg	Thr	Gly	Gly	Thr	Ile	Ser	Thr
			165					170					175		
Gly	Asn	His	Phe	Asp	Gly	Trp	Ala	Thr	Phe	Gly	Met	Asn	Leu	Gly	Thr
			180					185					190		
Phe	Asn	Tyr	Gln	Ile	Val	Ala	Thr	Glu	Gly	Tyr	Gln	Ser	Ser	Gly	Asn
		195				200						205			
Ser	Asp	Ile	Thr	Val	Ser	Asp	Gly	Gly	Ser	Ser	Ser	Ser	Ser	Ser	Ser
	210					215					220				
Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Gly	Gly	Gly	Gly	Thr	Lys	Ser
225					230					235					240
Phe	Thr	Val	Arg	Ala	Arg	Gly	Thr	Ala	Gly	Gly	Glu	Ser	Ile	Ser	Leu
			245					250					255		
Arg	Val	Asn	Asn	Thr	Asn	Val	Gln	Thr	Trp	Ser	Leu	Thr	Thr	Ser	Tyr
			260				265						270		
Gln	Asn	Leu	Thr	Ala	Ser	Thr	Thr	Leu	Thr	Gly	Gly	Ile	Thr	Val	Asn
		275					280					285			
Tyr	Thr	Asn	Asp	Ser	Ser	Gly	His	Asp	Val	Gln	Val	Asp	Tyr	Ile	Ile
	290					295					300				
Val	Asn	Gly	Gln	Thr	Arg	Gln	Ser	Glu	Ala	Gln	Ser	Tyr	Asn	Thr	Gly
305					310					315					320
Leu	Tyr	Ala	Asn	Gly	Arg	Cys	Gly	Gly	Gly	Tyr	Ser	Glu	Trp	Met	

His Cys Asn Gly 325 Ala Ile Gly Tyr Gly 330 Asn Thr Pro 335
 340 345

<210> 167
 <211> 669
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 167
 gtgaagctga aaagactggt caagatcgga ctgctgccgg ccgtattggt gtttagtgca 60
 acgcagcagt taaccgcgca aaccatctgc agcaaccaga ccggcaccaa caacggctac 120
 ttctactcgt tctggaagga caccgggtcg gcgtgcatga cactgggttc cggcggcaac 180
 tacagcgtca actggaacct gggttccggg aacatggtct gcggcaaagg ctggagtacc 240
 ggatcttcaa gccgcagaat cggctacaac gccggcgtct gggcgccgaa cggcaatgcc 300
 tacctgactc tgtatgggtg gaccaggaac ccgctcatcg agtactacgt ggtcgacagt 360
 tggggaagct ggaggccgcc aggcggaacc tccgcgggca ccgtcaatag cgatggcggg 420
 acctacaacc tctatcggac gcagcgggtc aacgcgcctt ccacgcgacg caccgggacg 480
 ttctatcagt actggagtgt ccggacctcg aagaggccca ccgggagcaa ccagaccatc 540
 accttcgcga accacgtgaa tgcgtggagg agcaaagggt ggaatctggg gagtcacgtc 600
 taccagataa tggcaacaga gggatatcaa agcagcggga attccaacct gacgggtgtg 660
 gcgcagtag 669

<210> 168
 <211> 222
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(36)

<400> 168
 Val Lys Leu Lys Arg Leu Phe Lys Ile Gly Leu Leu Pro Ala Val Leu
 1 5 10 15
 Leu Phe Ser Ala Thr Gln Gln Leu Thr Ala Gln Thr Ile Cys Ser Asn
 20 25 30
 Gln Thr Gly Thr Asn Asn Gly Tyr Phe Tyr Ser Phe Trp Lys Asp Thr
 35 40 45
 Gly Ser Ala Cys Met Thr Leu Gly Ser Gly Gly Asn Tyr Ser Val Asn
 50 55 60
 Trp Asn Leu Gly Ser Gly Asn Met Val Cys Gly Lys Gly Trp Ser Thr
 65 70 75 80
 Gly Ser Ser Ser Arg Arg Ile Gly Tyr Asn Ala Gly Val Trp Ala Pro
 85 90 95
 Asn Gly Asn Ala Tyr Leu Thr Leu Tyr Gly Trp Thr Arg Asn Pro Leu
 100 105 110
 Ile Glu Tyr Tyr Val Val Asp Ser Trp Gly Ser Trp Arg Pro Pro Gly
 115 120 125
 Gly Thr Ser Ala Gly Thr Val Asn Ser Asp Gly Gly Thr Tyr Asn Leu
 130 135 140
 Tyr Arg Thr Gln Arg Val Asn Ala Pro Ser Ile Asp Gly Thr Arg Thr
 145 150 155 160
 Phe Tyr Gln Tyr Trp Ser Val Arg Thr Ser Lys Arg Pro Thr Gly Ser
 165 170 175
 Asn Gln Thr Ile Thr Phe Ala Asn His Val Asn Ala Trp Arg Ser Lys
 180 185 190
 Gly Trp Asn Leu Gly Ser His Val Tyr Gln Ile Met Ala Thr Glu Gly
 195 200 205
 Tyr Gln Ser Ser Gly Asn Ser Asn Leu Thr Val Trp Ala Gln
 210 215 220

<210> 169
 <211> 1041

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 169

atgattgtta	gtttcaagag	cgtgaaggca	ctcgcgtgcc	tcgccgtgct	cggcattacc	60
gccgcgcagg	cgaaacctg	catcacttcc	agccagaccg	gtaccaacaa	cggcaactac	120
ttttccttct	ggaaggacag	cccgggtacc	gtcaacttct	gcatgtatgc	caatgggcgc	180
tacacctcca	actggagcgg	catcaacaac	tgggtgggcg	gcaagggctg	gcagacgggc	240
tccaaccgca	cggtgacctt	ctccggttcg	ttcaattcgc	ccggcaatgg	ctatctcacc	300
ttgtacggat	ggaccacgaa	tccattgatc	gagtactaca	tcgtcgacag	ctggggcacc	360
tatcgaccgc	cgggcggcca	gggcttcatg	ggcaccgtca	acagcgatgg	cggcacctat	420
gacatctacc	gcacgcagcg	cgtgaaccag	ccttccatca	tcggcaccgc	cacgttctac	480
cagtactgga	gcgtgcggca	gtcgaagcgc	gtcggcggca	cgatcaccac	ggccaaccac	540
ttcaacgcct	gggccacgct	gggcatgaac	ctggggcagc	acaactacca	ggcatggcc	600
accgagggtt	accagagcag	tggcagctcc	gacatcaccg	tgaccgaggg	cggcggctcc	660
tcgtcgtcca	gtggcggcgg	cagcaccagc	agtggcgggt	gcggcagcaa	gagcttcacc	720
gtgcgtgcgc	gcggcacggg	cggcggcgaa	aacatccagc	tgagggtcaa	caaccagacg	780
gtggcgaagt	gggaacctgac	caccagcatg	cagaactaca	acgcctcgac	cagcctgagt	840
ggcggcatca	ccgtcgtgta	caccaatgac	agcggcagcc	gcgacgtgca	ggtggactac	900
atcgctcgta	acggccagac	ccgccagtcc	gaagcccaga	gctacaacac	cgggctctat	960
gccaacggac	gttggtggtg	cggctcgaac	agcgagtgga	tgcatgcaa	cggcgcgatt	1020
ggctacggca	acacgcctta	g				1041

<210> 170

<211> 346

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(24)

<400> 170

Met	Ile	Val	Ser	Phe	Lys	Ser	Val	Lys	Ala	Leu	Ala	Cys	Leu	Ala	Val
1				5					10					15	
Leu	Gly	Ile	Thr	Ala	Ala	Gln	Ala	Gln	Thr	Cys	Ile	Thr	Ser	Ser	Gln
			20					25					30		
Thr	Gly	Thr	Asn	Asn	Gly	Asn	Tyr	Phe	Ser	Phe	Trp	Lys	Asp	Ser	Pro
		35				40						45			
Gly	Thr	Val	Asn	Phe	Cys	Met	Tyr	Ala	Asn	Gly	Arg	Tyr	Thr	Ser	Asn
	50					55					60				
Trp	Ser	Gly	Ile	Asn	Asn	Trp	Val	Gly	Gly	Lys	Gly	Trp	Gln	Thr	Gly
65				70						75				80	
Ser	Asn	Arg	Thr	Val	Thr	Tyr	Ser	Gly	Ser	Phe	Asn	Ser	Pro	Gly	Asn
			85					90						95	
Gly	Tyr	Leu	Thr	Leu	Tyr	Gly	Trp	Thr	Thr	Asn	Pro	Leu	Ile	Glu	Tyr
		100						105					110		
Tyr	Ile	Val	Asp	Ser	Trp	Gly	Thr	Tyr	Arg	Pro	Pro	Gly	Gly	Gln	Gly
	115					120						125			
Phe	Met	Gly	Thr	Val	Asn	Ser	Asp	Gly	Gly	Thr	Tyr	Asp	Ile	Tyr	Arg
	130					135					140				
Thr	Gln	Arg	Val	Asn	Gln	Pro	Ser	Ile	Ile	Gly	Thr	Ala	Thr	Phe	Tyr
145				150						155				160	
Gln	Tyr	Trp	Ser	Val	Arg	Gln	Ser	Lys	Arg	Val	Gly	Gly	Thr	Ile	Thr
			165					170						175	
Thr	Ala	Asn	His	Phe	Asn	Ala	Trp	Ala	Thr	Leu	Gly	Met	Asn	Leu	Gly
		180						185					190		
Gln	His	Asn	Tyr	Gln	Val	Met	Ala	Thr	Glu	Gly	Tyr	Gln	Ser	Ser	Gly
	195					200						205			
Ser	Ser	Asp	Ile	Thr	Val	Thr	Glu	Gly	Gly	Gly	Ser	Ser	Ser	Ser	Ser
	210					215					220				
Gly	Gly	Gly	Ser	Thr	Ser	Gly	Gly	Gly	Gly	Gly	Ser	Lys	Ser	Phe	Thr
225				230				235						240	
Val	Arg	Ala	Arg	Gly	Thr	Val	Gly	Gly	Glu	Asn	Ile	Gln	Leu	Gln	Val

245 250 255
 Asn Asn Gln Thr Val Ala Ser Trp Asn Leu Thr Thr Ser Met Gln Asn
 260 265 270
 Tyr Asn Ala Ser Thr Ser Leu Ser Gly Gly Ile Thr Val Val Tyr Thr
 275 280 285
 Asn Asp Ser Gly Ser Arg Asp Val Gln Val Asp Tyr Ile Val Val Asn
 290 295 300
 Gly Gln Thr Arg Gln Ser Glu Ala Gln Ser Tyr Asn Thr Gly Leu Tyr
 305 310 315 320
 Ala Asn Gly Arg Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His Cys
 325 330 335
 Asn Gly Ala Ile Gly Tyr Gly Asn Thr Pro
 340 345

<210> 171
 <211> 678
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 171
 atggagtga aaaaaatatac cagaaaagga ctgccactag tattcttgtc cttgttgttg 60
 ttcagtgtaa cgcagcagtc aaacgcccaa accatctgca gcaatcaaac tggcacaac 120
 aacggtttct tctattcggt ttggaaggac accggatcag catgcatgac tttgggctct 180
 ggcggcaatt acgacgtaag ttggaatctg ggttctggga atatggttgt cggcaaaggc 240
 tggagtaccg gatcatcaac caggagagta ggctacaatg ccggcatctg gcagccgaac 300
 ggcaatgcat atttggtctt ctatgggtgg acgagaaacc cacttataga atattacgtc 360
 gttgatagct ggggcacttt caggccgcct ggaggaacgt caataggctc cgtcaccact 420
 gatggtggtg cataccaaat atatcggacc cagcgagtca acgcgccttc cattgacggc 480
 gccagaactt tttatcagta ctggagtgtc cggacctcga agagaccgac cgggagcaac 540
 caaacatca cctttgcgaa tcacgttaac gcgtggagga atctagggtt gaatctgggg 600
 agtcatgttt accagataat ggccacagag ggatttcata gcagtgggag atctaacct 660
 acggtgtggt cacagtaa

<210> 172
 <211> 225
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(29)

<400> 172
 Met Glu Leu Lys Lys Ile Ser Arg Lys Gly Leu Pro Leu Val Phe Leu
 1 5 10 15
 Ser Leu Leu Leu Phe Ser Val Thr Gln Gln Ser Asn Ala Gln Thr Ile
 20 25 30
 Cys Ser Asn Gln Thr Gly Thr Asn Asn Gly Phe Phe Tyr Ser Phe Trp
 35 40 45
 Lys Asp Thr Gly Ser Ala Cys Met Thr Leu Gly Ser Gly Gly Asn Tyr
 50 55 60
 Asp Val Ser Trp Asn Leu Gly Ser Gly Asn Met Val Val Gly Lys Gly
 65 70 75 80
 Trp Ser Thr Gly Ser Ser Thr Arg Arg Val Gly Tyr Asn Ala Gly Ile
 85 90 95
 Trp Gln Pro Asn Gly Asn Ala Tyr Leu Ala Leu Tyr Gly Trp Thr Arg
 100 105 110
 Asn Pro Leu Ile Glu Tyr Tyr Val Val Asp Ser Trp Gly Thr Phe Arg
 115 120 125
 Pro Pro Gly Gly Thr Ser Ile Gly Ser Val Thr Thr Asp Gly Gly Thr
 130 135 140
 Tyr Gln Ile Tyr Arg Thr Gln Arg Val Asn Ala Pro Ser Ile Asp Gly
 145 150 155 160
 Ala Arg Thr Phe Tyr Gln Tyr Trp Ser Val Arg Thr Ser Lys Arg Pro

Thr Gly Ser Asn¹⁶⁵ Gln Thr Ile Thr Phe¹⁷⁰ Ala Asn His Val Asn¹⁷⁵ Ala Trp
 Arg Asn¹⁸⁰ Leu Gly Leu Asn Leu Gly¹⁸⁵ Ser His Val Tyr Gln¹⁹⁰ Ile Met Ala
 Thr Gln¹⁹⁵ Gly Phe His Ser Ser²⁰⁰ Gly Arg Ser Asn²⁰⁵ Leu Thr Val Trp Ser
 Gln²¹⁰
 225

<210> 173
 <211> 1503
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 173
 ttgaaaaaac tcgcagctgc cttatcactt gcaattacct ttgccgtacc gacaatagta 60
 caagcacaaag gtccacacatg gactaccagc acaatacaga aatacaacaa ctacgactat 120
 gaactctgga atgaaaacaa tcagggtacc gtttccatga agctcacagg agataacggt 180
 accgctgcca atgcggtagg cggaacgttt gagtctactt ggagtggtac aaagaatgtg 240
 cttttccgtt ccggcagaaa gtttacgggt acttcagggc aaagcgttga tgggtggcgg 300
 gctggcaaaa ccgctagtgc ttacggcaat ataagcatta acttcgccgc tacgtggtct 360
 tccggtgacg atgtgaagat gcttggcgta tatggttggg cgttttacgc actgccaaagt 420
 gtaccagaca aacaggaaaa cggcacttct actaattttt ccaatcaaat agaatactac 480
 atcattcaag accgcggcag ctataactcg gctacagggt gcaccaactc aaagaaatc 540
 ggtgaggcta ccattgacgg cattgcttat gagttccgtg tatgtgatag aatagggcaa 600
 cctatgttaa ctggcaacgg gaattttaag cagtatttca gtgttcctaa aagcactata 660
 aaccaccgca ccagcggtag aatctctgtt tccaaacact ttgaagaatg ggaaaaagtc 720
 ggcataaaaa tggacgggtc cttatacga gtagcgatga aagttgaatc ctattctggc 780
 aatgggaata gtaacggcaa tgctaaaatt acaagaata ttttgaccat tggcggaaca 840
 accacaactc aaagcagttc aagcggaggt tcaacgggtc cagatgaatg tggcgaatat 900
 aaaaagagtt tctgtggtgg cttgggatat ggaagcgtat attccaattt aaccgcaata 960
 ccctcaacgg gcgactgctt atacatcgga gattttgaag taatccagcc agctttgaat 1020
 tcaaccgttg ccataaacgg tgtggaat acctgcggaa gcgagtggtc agattgccct 1080
 tacaatgata aacccgattc aaaaaagat ggcggtctatt atgtttatgt gaaaacaggc 1140
 tcaattaaaca attatgagaa taacgggtgg caaaacattg tagctaaagc aaaaccggct 1200
 tgcacaccac cttctagcag ttccgggtgc gcaccaggtt cttcttcttc agacgaagaa 1260
 gaccagagc caattttgaa aaatcgcat cctataactc atttttccct tcaaacgctt 1320
 agcgataaag ccttgcgcat agaagtaaat gctccaacta ttgtggacat ttttgacctg 1380
 agagggaata aggttaaaag tttgaatgtt tacggttcgc aaaggggtta attatccctg 1440
 ccgagcgggg tgtattttgc caaagtgcgc gggatgaaaa gcgttagatt tgtgttgagg 1500
 taa 1503

<210> 174
 <211> 500
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(22)

<400> 174
 Leu Lys Lys Leu Ala Ala Ala Leu Ser Leu Ala Ile Thr Phe Ala Val
 1 5 10 15
 Pro Thr Ile Val Gln Ala Gln Gly Pro Thr Trp Thr Thr Ser Thr Ile
 20 25 30
 Gln Lys Tyr Asn Asn Tyr Asp Tyr Gln Leu Trp Asn Gln Asn Asn Gln
 35 40 45
 Gly Thr Val Ser Met Lys Leu Thr Gly Asp Asn Gly Thr Ala Ala Asn
 50 55 60
 Ala Val Gly Gly Thr Phe Gln Ser Thr Trp Ser Gly Thr Lys Asn Val
 65 70 75 80
 Leu Phe Arg Ser Gly Arg Lys Phe Thr Gly Thr Ser Gly Gln Ser Val

Asp Gly Gly Gly 85 Ala Gly Lys Thr Ala 90 Ser Ala Tyr Gly Asn 95 Ile Ser
 Ile Asn Phe Ala Ala Thr Trp Ser Ser Gly Asp Asp Val Lys Met Leu
 Gly Val 115 Tyr Gly Trp Ala Phe Tyr Ala Leu Pro Ser Val Pro Asp Lys
 Gln Glu Asn Gly Thr Ser Thr Asn Phe Ser Asn Gln Ile Glu Tyr Tyr
 145 Ile Ile Gln Asp Arg Gly Ser Tyr Asn Ser Ala Thr Gly Gly Thr Asn
 Ser Lys Lys Tyr 165 Gly Glu Ala Thr Ile Asp Gly Ile Ala Tyr Glu Phe
 Arg Val Cys Asp Arg Ile Gly Gln Pro Met Leu Thr Gly Asn Gly Asn
 Phe Lys 195 Gln Tyr Phe Ser Val 200 Pro Lys Ser Thr Ile Asn His Arg Thr
 Ser Gly Thr Ile Ser Val Ser Lys His Phe Glu Glu Trp Glu Lys Val
 225 Gly Met Lys Met Asp 230 Gly Pro Leu Tyr Glu Val Ala Met Lys Val Glu
 Ser Tyr Ser Gly Asn Gly Asn Ser Asn Gly Asn Ala Lys Ile Thr Lys
 Asn Ile Leu Thr Ile Gly Gly Thr Thr Thr Thr Gln Ser Ser Ser Ser
 Gly Gly 275 Ser Thr Val Pro Asp 280 Glu Cys Gly Glu Tyr 285 Lys Lys Ser Phe
 Cys Gly Gly Leu Gly Tyr Gly Ser Val Tyr Ser Asn Leu Thr Ala Ile
 305 Pro Ser Thr Gly Asp 310 Cys Leu Tyr Ile Gly Asp Phe Glu Val Ile Gln
 Pro Ala Leu Asn Ser Thr Val Ala Ile Asn Gly Val Glu Asn Thr Cys
 Gly Ser Glu Trp Ser Asp Cys Pro Tyr Val Lys Thr Gly Ser Ile Asn Asn
 Lys Asp Gly Gly Tyr Tyr Val Tyr Val Lys Thr Gly Ser Ile Asn Asn
 Tyr Glu Asn Asn Gly Trp Gln Asn Ile Val Ala Lys Ala Lys Pro Ala
 385 Cys Thr Pro Pro Ser Ser Ser Ser Gly Ala Ala Pro Gly Ser Ser Ser
 Ser Asp Glu Glu Asp Pro Glu Pro Ile Leu Lys Asn Arg Ile Pro Ile
 Thr His Phe 420 Ser Leu Gln Thr Leu Ser Asp Lys Ala Leu Arg Ile Glu
 Val Asn Ala Pro Thr Ile Val Asp Ile Phe Asp Leu Arg Gly Asn Lys
 Val Lys Ser Leu Asn Val Tyr Gly Ser Gln Arg Val Lys Leu Ser Leu
 465 Pro Ser Gly Val Tyr Phe Ala Lys Val Arg Gly Met Lys Ser Val Arg
 Phe Val Leu Arg 485
 500

<210> 175

<211> 1053

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 175

atgaagtcca	ttcgcagccg	cagcctcgcc	accgccgtcc	tggctggcgc	cctcggcgctc	60
gcagccgcag	gcgcgcaggc	gcagacgctc	aacaacaatt	ccaccggcac	gcacgacggc	120
tactactaca	cgttctggaa	ggactcgggc	agcgcctcga	tgaccctcca	tccgggaggga	180
cgctacagct	cccagtgagc	cagcaacacc	aacaactggg	tcggcgaggaa	aggctgggaat	240
cccgggtggc	cgcgctgggt	caactactcg	ggctactacg	gggtcaacaa	cagccagaac	300
tcctacctgg	cgctgtacgg	ctggacccgc	aatccgctgg	tcgagtacta	cgtgatcgag	360

agctacggct	cctacaaccc	ggccagttgc	gccggcgggg	tggaactacgg	cagctttccag	420
agcgatggcg	ccacctacaa	cgtacgtcgc	tgccctgcgcc	agaacgcgcc	gtcgatcgaa	480
ggcaacaaca	gcaccttcta	ccagtacttc	agcgtgcgca	atcccaagaa	gggattcggc	540
aacatctccg	gcacgatcac	cgtcgccaac	cacttcaact	actgggccag	ccgcggcctc	600
aacctcggca	accacgacta	catgggtgtc	gccaccgagg	gctaccagag	ccagggcagc	660
agcgacatca	ccgtgagttc	gggtaccggc	ggcggcggtg	gcggcgga	cacgggcagc	720
aagaccatcg	tggtgcgcgc	gcgcggcacc	gccggcgga	agaacatctc	gctcaaggtc	780
aacaacgcca	ccatcgccag	ctggacgctc	accaccagca	tgccaacta	cacggccacc	840
acctcggcat	cgggcggctc	gctgggtggag	ttcaccaacg	acggcgga	ccgcgacgtg	900
caggtggact	acctcagcgt	caatggcgcc	gtccgccagg	ccgaggacca	gacctacaac	960
accggcgtgt	accagaacgg	ccagtgcggc	ggcggcaacg	gccgcagcga	atggctgcac	1020
tgcaacggtg	ccatcggtct	cggaatctc	tga			1053

<210> 176

<211> 350

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(27)

<400> 176

Met	Lys	Ser	Ile	Arg	Ser	Arg	Ser	Leu	Ala	Thr	Ala	Val	Leu	Ala	Gly
1				5					10					15	
Ala	Leu	Gly	Val	Ala	Ala	Ala	Gly	Ala	Gln	Ala	Gln	Thr	Leu	Asn	Asn
			20					25					30		
Asn	Ser	Thr	Gly	Thr	His	Asp	Gly	Tyr	Tyr	Tyr	Thr	Phe	Trp	Lys	Asp
		35					40					45			
Ser	Gly	Ser	Ala	Ser	Met	Thr	Leu	His	Pro	Gly	Gly	Arg	Tyr	Ser	Ser
	50					55					60				
Gln	Trp	Thr	Ser	Asn	Thr	Asn	Asn	Trp	Val	Gly	Gly	Lys	Gly	Trp	Asn
65					70					75					80
Pro	Gly	Gly	Pro	Arg	Val	Val	Asn	Tyr	Ser	Gly	Tyr	Tyr	Gly	Val	Asn
				85					90					95	
Asn	Ser	Gln	Asn	Ser	Tyr	Leu	Ala	Leu	Tyr	Gly	Trp	Thr	Arg	Asn	Pro
			100					105					110		
Leu	Val	Glu	Tyr	Tyr	Val	Ile	Glu	Ser	Tyr	Gly	Ser	Tyr	Asn	Pro	Ala
		115					120					125			
Ser	Cys	Ala	Gly	Gly	Val	Asp	Tyr	Gly	Ser	Phe	Gln	Ser	Asp	Gly	Ala
	130					135					140				
Thr	Tyr	Asn	Val	Arg	Arg	Cys	Leu	Arg	Gln	Asn	Ala	Pro	Ser	Ile	Glu
145					150					155					160
Gly	Asn	Asn	Ser	Thr	Phe	Tyr	Gln	Tyr	Phe	Ser	Val	Arg	Asn	Pro	Lys
				165					170					175	
Lys	Gly	Phe	Gly	Asn	Ile	Ser	Gly	Thr	Ile	Thr	Val	Ala	Asn	His	Phe
			180					185					190		
Asn	Tyr	Trp	Ala	Ser	Arg	Gly	Leu	Asn	Leu	Gly	Asn	His	Asp	Tyr	Met
	195						200					205			
Val	Phe	Ala	Thr	Glu	Gly	Tyr	Gln	Ser	Gln	Gly	Ser	Ser	Asp	Ile	Thr
	210					215					220				
Val	Ser	Ser	Gly	Thr	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Asn	Thr	Gly	Ser
225					230					235					240
Lys	Thr	Ile	Val	Val	Arg	Ala	Arg	Gly	Thr	Ala	Gly	Gly	Glu	Asn	Ile
			245						250					255	
Ser	Leu	Lys	Val	Asn	Asn	Ala	Thr	Ile	Ala	Ser	Trp	Thr	Leu	Thr	Thr
			260					265					270		
Ser	Met	Ala	Asn	Tyr	Thr	Ala	Thr	Thr	Ser	Ala	Ser	Gly	Gly	Ser	Leu
	275						280					285			
Val	Glu	Phe	Thr	Asn	Asp	Gly	Gly	Asn	Arg	Asp	Val	Gln	Val	Asp	Tyr
	290					295					300				
Leu	Ser	Val	Asn	Gly	Ala	Val	Arg	Gln	Ala	Glu	Asp	Gln	Thr	Tyr	Asn
305					310					315					320
Thr	Gly	Val	Tyr	Gln	Asn	Gly	Gln	Cys	Gly	Gly	Gly	Asn	Gly	Arg	Ser
			325						330					335	
Glu	Trp	Leu	His	Cys	Asn	Gly	Ala	Ile	Gly	Phe	Gly	Asn	Leu		
			340					345					350		

<210> 177
 <211> 1299
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 177
 atgaaattgt tgaaaacgca caggcgtgcg attgctgccg cagcactagc ggtggcgact 60
 gttccaatcg ctcatgcgca aacgccttagc tcaaattgcca ctggaaccca gaatgggttac 120
 tactattcgt tttggaagga ttccggtaac gccaccatga cactcgggtgc cgggtggaac 180
 tattcttcat cctggaacag cagcactaac aactgggttg gcggtaaagg ctggatgccg 240
 ggtactcggc gcacagtcac ctattcgggc agttatagcg cgagtggaaac cagctacctc 300
 gcactttacg gctggactcg aaaccgctg atcgaatatt acattgtcga aaactgggtc 360
 aattacaatc ctgctgcccg cgcaacgaat tatgggactg tcaatattga cggcagcacc 420
 taccagctgg gccgcagcca acgggttaat cagccatcta ttgaaggcac ggccacgttc 480
 taccaatact ggagtgtgcg ccaaaacaag cgcaccagcg gaacgattaa tattggagcg 540
 catttcgatg catgggctgc tgtgggcttg aacctgggga ctacgatta tcagattatg 600
 gcgaccgagg gctaccagag cagcggccag tccaatatca cggtgagcga aggcagtagc 660
 ggcagcacga cttcgcagc atccagctcc agctcaagta cgagtccag tagttcttcc 720
 agcagttctt ccggcggcgg cacaggaagt tgtgcccggag tgaatgtgta cccaattgg 780
 accgcacgcg actggctctg cggcgcatac aatcacgccga atgcccgtga ccaaattggtc 840
 tatcaaaaca atttgtaccg ggcaaaactg tacaccaact ccacgcctgg aagcagtgcc 900
 tcctggacca gtctcgggtc ctgtagcggg ggcggtagca ccagttcaac aacgagctcc 960
 tccagttcct cttccacctc ggcgtcgcag agctccaact catccagcag cagttcaagc 1020
 agtccagca gcggtggctg tcgggaaatg tgtaactggt acggacaggg tatgtatcct 1080
 ctgtgtcaga acaccagcgg ttggggatgg gaaaataacc agaactgtat cggtcgccaa 1140
 acctgtcaaa gtcagaacgg cggctccggg ggtgtggtga acagctgtgg taccagcagc 1200
 tcttcgtcca gtagcacctc ctcatcgagc agttcaagtt cgctcagtggt caccacgtca 1260
 tcgtcctccg gaattcctgc agcccggggg atccactag 1299

<210> 178
 <211> 432
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(26)

<400> 178
 Met Lys Leu Leu Lys Thr His Arg Arg Ala Ile Ala Ala Ala Ala Leu
 1 5 10 15
 Ala Val Ala Thr Val Pro Ile Ala His Ala Gln Thr Leu Ser Ser Asn
 20 25 30
 Ala Thr Gly Thr Gln Asn Gly Tyr Tyr Tyr Ser Phe Trp Lys Asp Ser
 35 40 45
 Gly Asn Ala Thr Met Thr Leu Gly Ala Gly Gly Asn Tyr Ser Ser Ser
 50 55 60
 Trp Asn Ser Ser Thr Asn Asn Trp Val Gly Gly Lys Gly Trp Met Pro
 65 70 75 80
 Gly Thr Arg Arg Thr Val Thr Tyr Ser Gly Ser Tyr Ser Ala Ser Gly
 85 90 95
 Thr Ser Tyr Leu Ala Leu Tyr Gly Trp Thr Arg Asn Pro Leu Ile Glu
 100 105 110
 Tyr Tyr Ile Val Glu Asn Trp Val Asn Tyr Asn Pro Ala Ser Gly Ala
 115 120 125
 Thr Asn Tyr Gly Thr Val Asn Ile Asp Gly Ser Thr Tyr Gln Leu Gly
 130 135 140
 Arg Ser Gln Arg Val Asn Gln Pro Ser Ile Glu Gly Thr Ala Thr Phe
 145 150 155 160
 Tyr Gln Tyr Trp Ser Val Arg Gln Asn Lys Arg Thr Ser Gly Thr Ile
 165 170 175
 Asn Ile Gly Ala His Phe Asp Ala Trp Ala Ala Val Gly Leu Asn Leu
 180 185 190

Gly Thr His Asp Tyr Gln Ile Met Ala Thr Glu Gly Tyr Gln Ser Ser
 195 200 205
 Gly Gln Ser Asn Ile Thr Val Ser Glu Gly Ser Ser Gly Ser Thr Thr
 210 215 220
 Ser Ser Thr Ser Ser Ser Ser Ser Thr Ser Ser Ser Ser Ser
 225 230 235
 Ser Ser Ser Ser Gly Gly Gly Thr Gly Ser Cys Ala Gly Val Asn Val
 245 250 255
 Tyr Pro Asn Trp Thr Ala Arg Asp Trp Ser Gly Gly Ala Tyr Asn His
 260 265 270
 Ala Asn Ala Gly Asp Gln Met Val Tyr Gln Asn Asn Leu Tyr Arg Ala
 275 280 285
 Asn Trp Tyr Thr Asn Ser Thr Pro Gly Ser Asp Ala Ser Trp Thr Ser
 290 295 300
 Leu Gly Ser Cys Ser Gly Gly Gly Ser Thr Ser Thr Thr Ser Ser
 305 310 315
 Ser Ser Ser Ser Ser Thr Ser Ala Ser Ser Ser Ser Asn Ser Ser Ser
 325 330 335
 Ser Ser Ser Ser Ser Ser Ser Ser Gly Cys Arg Glu Met Cys Asn
 340 345 350
 Trp Tyr Gly Gln Gly Met Tyr Pro Leu Cys Gln Asn Thr Ser Gly Trp
 355 360 365
 Gly Trp Glu Asn Asn Gln Asn Cys Ile Gly Arg Gln Thr Cys Gln Ser
 370 375 380
 Gln Asn Gly Gly Ser Gly Gly Val Val Asn Ser Cys Gly Thr Ser Ser
 385 390 395
 Ser Ser Ser Ser Ser Thr Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
 405 410 415
 Gly Thr Thr Ser Ser Ser Gly Ile Pro Ala Ala Arg Gly Ile His
 420 425 430

<210> 179
 <211> 852
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 179
 atgaagaatt ggccgggaac ggggtattata ttattattgg cgggcggcct tttggcggct 60
 tgtttgacgg gcaaacggca agagggggcaa aaagtggatc cggatactca aaacgagaaa 120
 ttgacaggcg ggaccgtgtt tacagctaac agcaggggga acaggcccct ggaagggttcg 180
 ccttatgggtt acgaaatgtg gacgcagggc ggggaataata acaagcttgt ttgggttcggg 240
 ccggatcagg ggggaggggc ggctttcagg gcagaatgga acgagccgga tgattttttg 300
 ggacgactgg gtttctggtg gggaaacggc gggcaattta aagaatataa aaatatgtac 360
 gcggaatttc attacacaag gtcggggcgc ggcaccggcg gcagttattc ttatataggc 420
 atttacggct gggcgagaaa cccgaacgcc gcgaacgagg aagacagggt aatagaatac 480
 tatattgtgg acgactggtt cggaatcaa tggcagtcgg acgacacccc cattaccaca 540
 agaacaacag gaggtccgt attgggtacc attatagcgg acggcgcgtt ttacaacgtc 600
 gtcaggaatg tgagaaccca aaagccttcg atagacggca tcaaacatt cgccaatac 660
 ttcagcatac gccaaacacc gcgcaaagc gggacaatct ccatcaccga acatttcaaa 720
 caatgggaaa gcatgggcct gaagctcggg aatatgtacg aggcaaaatt cctggtagaa 780
 gccggcggcg gcaccggctg gctggagttt acgtatctta aactgacgca ggaagaaaaa 840
 aaaagaatt ag 852

<210> 180
 <211> 283
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(19)

<400> 180
 Met Lys Asn Trp Pro Gly Thr Gly Ile Ile Leu Leu Leu Ala Gly Gly
 Page 139

<210>	182
<211>	358
<212>	PRT
<213>	Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(25)

<400> 182

```

Met Asn Phe Ser Leu Arg Lys Ala Ala Ala Ala Leu Ala Cys Val Ala
 1      5      10      15
Gly Leu Tyr Ala Ser Ser Ala Gly Ala Gln Thr Cys Leu Thr Asn Asn
 20      25      30
Gln Thr Gly Asn Asn Gly Gly Tyr Tyr Ser Phe Trp Lys Asp Ser
 35      40      45
Gly Asn Val Thr Phe Cys Leu Gln Ser Gly Gly Arg Tyr Thr Ser Gln
 50      55      60
Trp Ser Asn Val Asn Asn Trp Val Gly Gly Lys Gly Trp Asn Pro Gly
 65      70      75      80
Gly Arg Arg Thr Val Thr Tyr Ser Gly Thr Tyr Asn Pro Asn Gly Asn
 85      90      95
Ser Tyr Leu Thr Leu Tyr Gly Trp Thr Thr Asn Pro Leu Val Glu Tyr
100      105      110
Tyr Ile Val Asp Ser Trp Gly Ser Trp Arg Pro Pro Gly Ser Gly Tyr
115      120      125
Met Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Thr
130      135      140
Gln Arg Val Asn Gln Pro Ser Ile Ile Gly Thr Ala Thr Phe Tyr Gln
145      150      155      160
Tyr Trp Ser Val Arg Gln Ser Lys Arg Val Gly Gly Thr Ile Thr Ser
165      170      175
Gly Asn His Phe Asp Ala Trp Ala Ser Leu Gly Met Asn Leu Gly Thr
180      185      190
His Asn Tyr Met Val Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly Ser
195      200      205
Ser Asp Ile Thr Val Gly Ser Gly Ser Ser Ser Ser Ser Ser Ser Ser
210      215      220
Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
225      230      235      240
Ser Ser Ser Gly Gly Gly Thr Lys Ser Phe Thr Val Arg Ala Arg
245      250      255
Gly Thr Ala Gly Gly Glu Ser Ile Thr Leu Arg Val Asn Asn Gln Asn
260      265      270
Val Gln Thr Trp Thr Leu Gly Thr Ser Met Gln Asn Tyr Thr Ala Ser
275      280      285
Thr Ser Leu Ser Gly Gly Ile Thr Val Ala Phe Thr Asn Asp Gly Gly
290      295      300
Asn Arg Asp Val Gln Val Asp Tyr Ile Ile Val Asn Gly Gln Thr Arg
305      310      315      320
Gln Ser Glu Ala Gln Thr Tyr Asn Thr Gly Leu Tyr Ala Asn Gly Arg
325      330      335
Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His Cys Asn Gly Ala Ile
340      345      350
Gly Tyr Gly Asn Thr Pro
355

```

<210> 183

<211> 1083

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 183

```

atgatcgaag gtctcaggag acctgccttc agtggcagga gcatcgtcaa ggcattgctc      60
tgcgtcgcgg ccctgtatgc atcggcggcg caggcgcaga cctgtctcag ttcgagccag      120
accggcacca acaacggctt ctactattcg ttctggaagg acagcccggg cagcgtgcag      180
ttctgcatgt attccggcgg ccgctacaca tccaactgga gcggcatcaa caactgggtc      240
ggcggcaagg ggtggcagac cggcgccctc cgctgggtca gctactcggg cacgttcaat      300
tcaccgggca acggctacct ggcgctgtac ggctggacca ccaatccact ggtcgagtac      360

```


tacatcgtcg	acaactgggg	cacctatcgc	ccgccggg	gcacgggatt	ccagggcacg	420
gtgaccagt	acggcggtac	ctacgacatc	taccggaccg	agcgaccaa	cgcgccctgc	480
atcaccggca	acaactgcaa	cttctcgcag	ttctggagcg	tgccgcagtc	gaagcgacc	540
ggcggcacca	tcaccaccgg	caatcacttc	agcgccctgg	cgtcgcacgg	catgaacatg	600
ggccagcaca	actaccagat	catggccacc	gaggggttacc	agagcaacgg	cagctcggac	660
atcacggtct	cgaagggcag	cagttcgtcg	agcagcagca	gttcgtcctc	ttcgtcgagc	720
agcagctcgt	cgagcggcgg	cggcggcagc	aagagcttca	cggtgcgcgc	ccgcggcacc	780
gcgggtggcg	agcagatccg	gctgcgcgtg	aacaatacga	ccgtgcagac	ctggacgctg	840
aacaccacga	tgacgaacta	caccgcttcg	accacgctga	gcggcgccat	cacgggtggag	900
tacttcaacg	acagcaccaa	tcacgacgtg	caggtggact	acatcatcgt	gaacggcgcg	960
acgcgccagt	ccgaagcgca	gagctacaac	accggcctgt	atgccaacgg	ccgttgcggt	1020
ggcggttcca	acagcgaatg	gatgcattgc	aatggcgcca	tcggctacgg	caacactcca	1080
taa						1083

<210> 184
 <211> 360
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(32)

<400> 184
 Met Ile Glu Gly Leu Arg Arg Pro Ala Phe Ser Gly Arg Ser Ile Val
 1 5 10 15
 Lys Ala Leu Leu Cys Val Ala Ala Leu Tyr Ala Ser Ala Ala Gln Ala
 20 25 30
 Gln Thr Cys Leu Ser Ser Ser Gln Thr Gly Thr Asn Asn Gly Phe Tyr
 35 40 45
 Tyr Ser Phe Trp Lys Asp Ser Pro Gly Ser Val Gln Phe Cys Met Tyr
 50 55 60
 Ser Gly Gly Arg Tyr Thr Ser Asn Trp Ser Gly Ile Asn Asn Trp Val
 65 70 75 80
 Gly Gly Lys Gly Trp Gln Thr Gly Ala Ser Arg Val Val Ser Tyr Ser
 85 90 95
 Gly Thr Phe Asn Ser Pro Gly Asn Gly Tyr Leu Ala Leu Tyr Gly Trp
 100 105 110
 Thr Thr Asn Pro Leu Val Glu Tyr Tyr Ile Val Asp Asn Trp Gly Thr
 115 120 125
 Tyr Arg Pro Pro Gly Gly Thr Gly Phe Gln Gly Thr Val Thr Ser Asp
 130 135 140
 Gly Gly Thr Tyr Asp Ile Tyr Arg Thr Glu Arg Thr Asn Ala Pro Cys
 145 150 155 160
 Ile Thr Gly Asn Asn Cys Asn Phe Ser Gln Phe Trp Ser Val Arg Gln
 165 170 175
 Ser Lys Arg Thr Gly Gly Thr Ile Thr Thr Gly Asn His Phe Ser Ala
 180 185 190
 Trp Ala Ser His Gly Met Asn Met Gly Gln His Asn Tyr Gln Ile Met
 195 200 205
 Ala Thr Glu Gly Tyr Gln Ser Asn Gly Ser Ser Asp Ile Thr Val Ser
 210 215 220
 Glu Gly Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
 225 230 235 240
 Ser Ser Ser Ser Ser Gly Gly Gly Ser Lys Ser Phe Thr Val Arg
 245 250 255
 Ala Arg Gly Thr Ala Gly Gly Glu Gln Ile Arg Leu Arg Val Asn Asn
 260 265 270
 Thr Thr Val Gln Thr Trp Thr Leu Asn Thr Thr Met Thr Asn Tyr Thr
 275 280 285
 Ala Ser Thr Thr Leu Ser Gly Gly Ile Thr Val Glu Tyr Phe Asn Asp
 290 295 300
 Ser Thr Asn His Asp Val Gln Val Asp Tyr Ile Ile Val Asn Gly Ala
 305 310 315 320
 Thr Arg Gln Ser Glu Ala Gln Ser Tyr Asn Thr Gly Leu Tyr Ala Asn
 325 330 335
 Gly Arg Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His Cys Asn Gly

Ala Ile Gly Tyr Gly Asn Thr Pro
 340 355 360 345 350

<210> 185
 <211> 684
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 185
 atgaatttga aaagattgag gctgttgttt gtgatgtgta ttggatttgt gctgacactg 60
 acggctgtgc cagctcatgc ggaaacgatt tatgataata ggatagggac acacagcggg 120
 tacgattttg aattatggaa ggattacgga aatacctcga tgacactcaa taacggcggg 180
 gcatttagtg caagctggaa caatattgga aatgccttat ttcgaaaagg aaagaagttt 240
 gattccacta aaactcatca tcaacttggc aacatctcca tcaactacaa cgcagccttt 300
 aacccgggcg ggaattccta tttatgtgtc tatggctgga cacaatctcc attagctgaa 360
 tactacattg ttgagtcattg gggcacatat cgtccaacag gaacgtataa aggatcattt 420
 tatgccgatg gaggcacata tgacatatat gaaacgctcc gtgtcaatca gccttctatc 480
 attggagacg ctaccttcaa acaatattgg agtgtacgtc aaacaaaacg cacaagcggg 540
 actgtttccg tcagttagca ttttaaaaaa tgggaaagct taggcatgcc aatgggaaaa 600
 atgtatgaaa cagcattaac tgtagaaggc taccgaagca acggaagtgc gaatgtcatg 660
 acgaatcagc tgatgattcg ataa 684

<210> 186
 <211> 227
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(27)

<400> 186
 Met Asn Leu Lys Arg Leu Arg Leu Leu Phe Val Met Cys Ile Gly Phe
 1 5 10 15
 Val Leu Thr Leu Thr Ala Val Pro Ala His Ala Glu Thr Ile Tyr Asp
 20 25 30
 Asn Arg Ile Gly Thr His Ser Gly Tyr Asp Phe Glu Leu Trp Lys Asp
 35 40 45
 Tyr Gly Asn Thr Ser Met Thr Leu Asn Asn Gly Gly Ala Phe Ser Ala
 50 55 60
 Ser Trp Asn Asn Ile Gly Asn Ala Leu Phe Arg Lys Gly Lys Lys Phe
 65 70 75 80
 Asp Ser Thr Lys Thr His His Gln Leu Gly Asn Ile Ser Ile Asn Tyr
 85 90 95
 Asn Ala Ala Phe Asn Pro Gly Gly Asn Ser Tyr Leu Cys Val Tyr Gly
 100 105 110
 Trp Thr Gln Ser Pro Leu Ala Glu Tyr Tyr Ile Val Glu Ser Trp Gly
 115 120 125
 Thr Tyr Arg Pro Thr Gly Thr Tyr Lys Gly Ser Phe Tyr Ala Asp Gly
 130 135 140
 Gly Thr Tyr Asp Ile Tyr Glu Thr Leu Arg Val Asn Gln Pro Ser Ile
 145 150 155 160
 Ile Gly Asp Ala Thr Phe Lys Gln Tyr Trp Ser Val Arg Gln Thr Lys
 165 170 175
 Arg Thr Ser Gly Thr Val Ser Val Ser Glu His Phe Lys Lys Trp Glu
 180 185 190
 Ser Leu Gly Met Pro Met Gly Lys Met Tyr Glu Thr Ala Leu Thr Val
 195 200 205
 Glu Gly Tyr Arg Ser Asn Gly Ser Ala Asn Val Met Thr Asn Gln Leu
 210 215 220
 Met Ile Arg
 225

<210> 187
 <211> 642
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 187
 atgtttaagt ttaaaaagaa tttcttagtt ggattatcgg cagctttaat gattattagc 60
 ttgttttcgg caaccgcctc tgcagctagc acagactact ggcaaaattg gactgatggg 120
 ggcggtatag taaacgctgt caatgggtct ggcgggaatt acagtgttaa ttggctaat 180
 accggaaatt ttgttggtgg taaagggttg actacaggtt cgccatttag gacgataaac 240
 tataatgccg gagtttgggc gccgaatggc aatggatatt taactttata tggttggacg 300
 agatcacctc tcatagaata ttatgtagtg gattcatggg gtacttatag acctactgga 360
 acgtataaag gtactgtaaa aagtgatggg ggtacatatg acatatatac aactacacgt 420
 tataacgcac cttccattga tggcgatcgc actactttta cgcagtactg gactgttcgc 480
 cagtcgaaga gaccaaccgg aagcaacgct acaatcactt tcagcaatca tgtgaacgca 540
 tggaagagcc atggaatgaa tctgggcagt aattgggctt accaagtcac ggcgacagaa 600
 ggatatcaaa gtagtggaag ttctaacgta acagtgtggt aa 642

<210> 188
 <211> 213
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 188
 Met Phe Lys Phe Lys Lys Asn Phe Leu Val Gly Leu Ser Ala Ala Leu
 1 5 10 15
 Met Ser Ile Ser Leu Phe Ser Ala Thr Ala Ser Ala Ala Ser Thr Asp
 20 25 30
 Tyr Trp Gln Asn Trp Thr Asp Gly Gly Ile Val Asn Ala Val Asn
 35 40 45
 Gly Ser Gly Gly Asn Tyr Ser Val Asn Trp Ser Asn Thr Gly Asn Phe
 50 55 60
 Val Val Gly Lys Gly Trp Thr Thr Gly Ser Pro Phe Arg Thr Ile Asn
 65 70 75 80
 Tyr Asn Ala Gly Val Trp Ala Pro Asn Gly Asn Gly Tyr Leu Thr Leu
 85 90 95
 Tyr Gly Trp Thr Arg Ser Pro Leu Ile Glu Tyr Tyr Val Val Asp Ser
 100 105 110
 Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys Gly Thr Val Lys Ser
 115 120 125
 Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr Thr Arg Tyr Asn Ala Pro
 130 135 140
 Ser Ile Asp Gly Asp Arg Thr Thr Phe Thr Gln Tyr Trp Ser Val Arg
 145 150 155 160
 Gln Ser Lys Arg Pro Thr Gly Ser Asn Ala Thr Ile Thr Phe Ser Asn
 165 170 175
 His Val Asn Ala Trp Lys Ser His Gly Met Asn Leu Gly Ser Asn Trp
 180 185 190
 Ala Tyr Gln Val Met Ala Thr Glu Gly Tyr Gln Ser Ser Ser Ser
 195 200 205
 Asn Val Thr Val Trp
 210

<210> 189
 <211> 570
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample
 Page 144

```

<400> 189
atggccctta tggcttcgac agactactgg caaaattgga ctgatggtgg tgggacagta      60
aatgctacca atggatctga tggcaattac agcgtttcat ggtcaaattg cgggaatttt      120
gttgttggtg aaggctggac taccggatca gcaactaggg taataaacta taatgccgga      180
gccttttcgc cgtccggtaa tggatatttg gctctttatg ggtggacgag aaattcactc      240
atagaatatt acgtcgttga tagctggggg acttatagac ctactggaac ttataaaggc      300
actgtgacta gtgatggagg gacttatgac atatacacga ctacacgaac caacgcacct      360
tccattgacg gcaataatac aactttcacc cagttctgga gtgttaggca gtcgaagaga      420
ccgattggta ccaacaatac catcaccttt agcaaccatg ttaacgcctg gaagagtaaa      480
ggaatgaatt tggggagtag ttggtcttat caggtattag caacagaggg ctatcaaagt      540
agtgggtact ctaacgtaac ggtctggtaa                                570

```

```

<210> 190
<211> 189
<212> PRT
<213> Unknown

```

```

<220>
<223> obtained from an environmental sample

```

```

<400> 190
Met Ala Leu Met Ala Ser Thr Asp Tyr Trp Gln Asn Trp Thr Asp Gly
 1      5      10      15
Gly Gly Thr Val Asn Ala Thr Asn Gly Ser Asp Gly Asn Tyr Ser Val
 20     25     30
Ser Trp Ser Asn Cys Gly Asn Phe Val Val Gly Lys Gly Trp Thr Thr
 35     40     45
Gly Ser Ala Thr Arg Val Ile Asn Tyr Asn Ala Gly Ala Phe Ser Pro
 50     55     60
Ser Gly Asn Gly Tyr Leu Ala Leu Tyr Gly Trp Thr Arg Asn Ser Leu
 65     70     75     80
Ile Glu Tyr Tyr Val Asp Ser Trp Gly Thr Tyr Arg Pro Thr Gly
 85     90     95
Thr Tyr Lys Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr
100    105    110
Thr Thr Thr Arg Thr Asn Ala Pro Ser Ile Asp Gly Asn Asn Thr Thr
115    120    125
Phe Thr Gln Phe Trp Ser Val Arg Gln Ser Lys Arg Pro Ile Gly Thr
130    135    140
Asn Asn Thr Ile Thr Phe Ser Asn His Val Asn Ala Trp Lys Ser Lys
145    150    155    160
Gly Met Asn Leu Gly Ser Ser Trp Ser Tyr Gln Val Leu Ala Thr Glu
165    170    175
Gly Tyr Gln Ser Ser Gly Tyr Ser Asn Val Thr Val Trp
180    185

```

```

<210> 191
<211> 1053
<212> DNA
<213> Unknown

```

```

<220>
<223> obtained from an environmental sample

```

```

<400> 191
atgaagtcca ttcgcagccg cagcctcgcc accgcgctcc tggctggcgc cctcggcgtc      60
gcagccgccc gcgcgcaggc gcagacgctc aacaacaatt ccaccggcac gcacgacggc      120
ttctactaca cgttctggaa ggactcgggc agcgctctga tgaccctcca tccgggcgga      180
cgctacagct cccagtggac cagcaacacc aacaactggg tcggcgggaa aggctggaat      240
cccgttgccc cgcgcgtggg caactactcg ggtactacg ggttcaacaa cagccagaac      300
tcctactctg cgctgtacgg ctggaccggc aatccgctgg tcgagtacta cgtgatcgag      360
agctacggct cctacaaccc ggccagttgc gccggcgggg tggactacgg cagcttcag      420
agcgatggcg ccacctacaa cgtacgcgcg tgcctgcgcc agaacgcgcc gtcgatcgaa      480
ggcaacaaca gcaccttcta ccagtacttc agcgtgcgca atcccaagaa gggattcggc      540
aacatctccg cgcagatcac cgtcgcaact catttcaact actgggccag ccgcggcctc      600
aacctcggca accacgacta catggtgttc gccaccgagg gctaccagag ccagggcagc      660
agcgacatca ccgtgagttc gggtaaccgg ggcggcgggt gcggcggcaa cacgggcagc      720
aagaccatcg tgggtgcgcgc gcgcggcacc gccggcggag agaacatctc gctcaaggtc      780

```

aacaacgcc	ccatcgccag	ctggacgctc	accaccagca	tggccaacta	cacggccacc	840
acctcggcat	cgggcggctc	gctgggtggag	ttcaccaacg	acggcggcaa	ccgcgacgtg	900
caggtggact	acctcagcgt	caatggcgcc	gtccgccagg	ccgaggacca	gacctacaac	960
accggcgtgt	accagaacgg	ccagtgcggc	ggcggcaacg	gccgcagcga	atggctgcac	1020
tgcaacggtg	ccatcggtctt	cggaaatctc	tga			1053

<210> 192
 <211> 350
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(27)

<400> 192

Met	Lys	Ser	Ile	Arg	Ser	Arg	Ser	Leu	Ala	Thr	Ala	Val	Leu	Ala	Gly
1				5					10				15		
Ala	Leu	Gly	Val	Ala	Ala	Ala	Gly	Ala	Gln	Ala	Gln	Thr	Leu	Asn	Asn
			20					25					30		
Asn	Ser	Thr	Gly	Thr	His	Asp	Gly	Phe	Tyr	Tyr	Thr	Phe	Trp	Lys	Asp
			35				40					45			
Ser	Gly	Ser	Ala	Ser	Met	Thr	Leu	His	Pro	Gly	Gly	Arg	Tyr	Ser	Ser
			50			55					60				
Gln	Trp	Thr	Ser	Asn	Thr	Asn	Asn	Trp	Val	Gly	Gly	Lys	Gly	Trp	Asn
65					70					75				80	
Pro	Gly	Gly	Pro	Arg	Val	Val	Asn	Tyr	Ser	Gly	Tyr	Tyr	Gly	Val	Asn
				85					90					95	
Asn	Ser	Gln	Asn	Ser	Tyr	Leu	Ala	Leu	Tyr	Gly	Trp	Thr	Arg	Asn	Pro
			100				105						110		
Leu	Val	Glu	Tyr	Tyr	Val	Ile	Glu	Ser	Tyr	Gly	Ser	Tyr	Asn	Pro	Ala
		115					120					125			
Ser	Cys	Ala	Gly	Gly	Val	Asp	Tyr	Gly	Ser	Phe	Gln	Ser	Asp	Gly	Ala
		130				135					140				
Thr	Tyr	Asn	Val	Arg	Arg	Cys	Leu	Arg	Gln	Asn	Ala	Pro	Ser	Ile	Glu
145					150					155				160	
Gly	Asn	Asn	Ser	Thr	Phe	Tyr	Gln	Tyr	Phe	Ser	Val	Arg	Asn	Pro	Lys
				165					170					175	
Lys	Gly	Phe	Gly	Asn	Ile	Ser	Gly	Thr	Ile	Thr	Val	Ala	Asn	His	Phe
			180				185						190		
Asn	Tyr	Trp	Ala	Ser	Arg	Gly	Leu	Asn	Leu	Gly	Asn	His	Asp	Tyr	Met
		195					200					205			
Val	Phe	Ala	Thr	Glu	Gly	Tyr	Gln	Ser	Gln	Gly	Ser	Ser	Asp	Ile	Thr
		210				215					220				
Val	Ser	Ser	Gly	Thr	Gly	Gly	Gly	Gly	Gly	Gly	Asn	Thr	Gly	Ser	
225					230					235				240	
Lys	Thr	Ile	Val	Val	Arg	Ala	Arg	Gly	Thr	Ala	Gly	Gly	Glu	Asn	Ile
				245					250					255	
Ser	Leu	Lys	Val	Asn	Asn	Ala	Thr	Ile	Ala	Ser	Trp	Thr	Leu	Thr	Thr
			260				265						270		
Ser	Met	Ala	Asn	Tyr	Thr	Ala	Thr	Thr	Ser	Ala	Ser	Gly	Gly	Ser	Leu
		275				280						285			
Val	Glu	Phe	Thr	Asn	Asp	Gly	Gly	Asn	Arg	Asp	Val	Gln	Val	Asp	Tyr
		290				295					300				
Leu	Ser	Val	Asn	Gly	Ala	Val	Arg	Gln	Ala	Glu	Asp	Gln	Thr	Tyr	Asn
305					310					315					320
Thr	Gly	Val	Tyr	Gln	Asn	Gly	Gln	Cys	Gly	Gly	Gly	Asn	Gly	Arg	Ser
				325					330					335	
Glu	Trp	Leu	His	Cys	Asn	Gly	Ala	Ile	Gly	Phe	Gly	Asn	Leu		
			340					345					350		

<210> 193
 <211> 840
 <212> DNA
 <213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 193

atgacgaagt	atcggttagg	aataggtatt	ttcattttgt	tggtttgttg	cttttcggcg	60
gcatgtattg	tcctaaaca	acaagaggaa	caaaaagtgg	ctcctacaga	attgaccggc	120
gcgataacat	tcacagccaa	cagcaacgga	aacaagcccc	tgaacggctc	gccctacggg	180
tacgaaatat	ggacacaggg	cgggaccaat	aacaaactga	tctggttcgg	gccggatcag	240
ggcggcggcg	cggctttcag	agccgaatgg	aacaacccta	acgatttttt	aggccgcgtg	300
ggtttttact	ggggtaatgg	cggaaaatat	accgagtaca	aaaatatgta	tgcggtattt	360
agctacacta	gatctggacg	caacaccgcc	ggtaattatt	catatatagg	gatttatggc	420
tgggctagaa	atccaaatgc	cgcaaaagaa	gaagacaaat	tgatagagta	ttatatgtg	480
gaagattggg	ttggcaatca	atggcaagag	gatagctcac	ccattaccac	taatacaaca	540
agtggaaaccg	tattgggaag	ttttactata	gatggcgcgg	tttataatgt	cgttagaaat	600
gtcagagtcc	aacaaccttc	gatagacgga	accaaacaat	tcaccaata	cttcagcata	660
cgacaaacgc	cccacagag	cgggacaatt	tccattaccg	ggcatttcag	gcaatgggag	720
agcatgggtt	tacagcttgg	caatatgtac	gaggcaaatg	ttcttggtga	agccggcggc	780
ggcacaggat	ggctggaatt	ttcatacctt	aaattaacga	tggaagacag	cttaaggtaa	840

<210> 194

<211> 279

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(21)

<400> 194

Met	Thr	Lys	Tyr	Arg	Leu	Gly	Ile	Gly	Ile	Phe	Ile	Leu	Leu	Val	Cys
1				5					10					15	
Cys	Phe	Ser	Ala	Ala	Cys	Ile	Val	Pro	Lys	Gln	Gln	Glu	Glu	Gln	Lys
			20					25				30			
Val	Ala	Pro	Thr	Glu	Leu	Thr	Gly	Ala	Ile	Thr	Phe	Thr	Ala	Asn	Ser
		35					40				45				
Asn	Gly	Asn	Lys	Pro	Leu	Asn	Gly	Ser	Pro	Tyr	Gly	Tyr	Glu	Ile	Trp
	50					55				60					
Thr	Gln	Gly	Gly	Thr	Asn	Asn	Lys	Leu	Ile	Trp	Phe	Gly	Pro	Asp	Gln
65					70				75					80	
Gly	Gly	Gly	Ala	Ala	Phe	Arg	Ala	Glu	Trp	Asn	Asn	Pro	Asn	Asp	Phe
			85					90					95		
Leu	Gly	Arg	Val	Gly	Phe	Tyr	Trp	Gly	Asn	Gly	Gly	Lys	Tyr	Thr	Glu
		100						105				110			
Tyr	Lys	Asn	Met	Tyr	Ala	Asp	Phe	Ser	Tyr	Thr	Arg	Ser	Gly	Arg	Asn
	115					120					125				
Thr	Ala	Gly	Asn	Tyr	Ser	Tyr	Ile	Gly	Ile	Tyr	Gly	Trp	Ala	Arg	Asn
	130				135					140					
Pro	Asn	Ala	Ala	Lys	Glu	Asp	Lys	Leu	Ile	Glu	Tyr	Tyr	Ile	Val	
145					150				155					160	
Glu	Asp	Trp	Phe	Gly	Asn	Gln	Trp	Gln	Glu	Asp	Ser	Ser	Pro	Ile	Thr
			165					170					175		
Thr	Asn	Thr	Thr	Ser	Gly	Thr	Val	Leu	Gly	Ser	Phe	Thr	Ile	Asp	Gly
		180						185					190		
Ala	Val	Tyr	Asn	Val	Val	Arg	Asn	Val	Arg	Val	Gln	Gln	Pro	Ser	Ile
	195						200				205				
Asp	Gly	Thr	Lys	Thr	Phe	Thr	Gln	Tyr	Phe	Ser	Ile	Arg	Gln	Thr	Pro
	210				215						220				
Arg	Gln	Ser	Gly	Thr	Ile	Ser	Ile	Thr	Gly	His	Phe	Arg	Gln	Trp	Glu
225					230					235				240	
Ser	Met	Gly	Leu	Gln	Leu	Gly	Asn	Met	Tyr	Glu	Ala	Lys	Phe	Leu	Val
			245						250				255		
Glu	Ala	Gly	Gly	Thr	Gly	Trp	Leu	Glu	Phe	Ser	Tyr	Leu	Lys	Leu	
		260					265					270			
Thr	Met	Glu	Asp	Ser	Leu	Arg									
		275													

<210> 195

<211> 1044

<212> DNA
<213> Unknown

<220>
<223> obtained from an environmental sample

<400> 195
atgttcaatc tgaagagagt ggcggcgctc ctgtgcgtcg cagggctggg ggtgtctgcg 60
gcaaattgcgc agacctgtct caattcgagt gggaccggca ccaacaacgg cttctattat 120
tccttctgga aagacagttc gggttcagtg aatttctgca tgtactccgg cggtcgctac 180
acgtcgagct ggagcggcat caacaactgg gtcggcggca agggctggca aaccggatcg 240
cgccggacca tcaactactc cggcagcttc aactcgccgg gcaatggcta cctcgcgctc 300
tacggatgga ccaccaatcc actcgtcgag tactacatcg tcgacaactg gggcacgtat 360
cgtccgcccg gcggccaggg ctacatgggc acggtcacga gcgacggcgc cacgtacgac 420
gtctatcgaa cgcaacgagt cgatgcgccc tcgatcattg gtgatcacca gaccttctat 480
caatactgga gcgtgcgtca gtcgaagagg accggcggaa ccatcaccac cggcaaccac 540
ttcgtggct ggcgcgagcta cggcatgaac ctggggaactc acaactacca gatcctggcg 600
accgaggggt atcaaagcag cggcagctcg gacctcaccg tgagcgaagg cagcagcagt 660
agcagcagcg gtggcgggag cagttcgagc agcagcggcg gcggtggcac caagagcttc 720
acggtccgcg cgcgcggcac ggccggtgga gagtcgatca cgttgcgcgt gaataaccag 780
aacgtgcaga cctggacgct cggcacgagc atgacgaact acacggcgtc gacgtcgctg 840
agcggcggca tcacgtggc gttcacgaac gacggtggca accgcgatgt tcaggtggac 900
tacatcatcg tgaacggcca gacacgccag tcggaagcgc agagctacaa caccgggctc 960
tacgcgaatg gacgttgccg cggtggtcgc aacagcgagt ggatgcactg caacggcgcg 1020
attggctacg gaaacacgcc gtaa 1044

<210> 196
<211> 347
<212> PRT
<213> Unknown

<220>
<223> obtained from an environmental sample

<221> SIGNAL
<222> (1)...(23)

<400> 196
Met Phe Asn Leu Lys Arg Val Ala Ala Leu Leu Cys Val Ala Gly Leu
1 5 10 15
Gly Val Ser Ala Ala Asn Ala Gln Thr Cys Leu Asn Ser Ser Gly Thr
20 25 30
Gly Thr Asn Asn Gly Phe Tyr Tyr Ser Phe Trp Lys Asp Ser Pro Gly
35 40 45
Ser Val Asn Phe Cys Met Tyr Ser Gly Gly Arg Tyr Thr Ser Ser Trp
50 55 60
Ser Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Gln Thr Gly Ser
65 70 75 80
Arg Arg Thr Ile Asn Tyr Ser Gly Ser Phe Asn Ser Pro Gly Asn Gly
85 90 95
Tyr Leu Ala Leu Tyr Gly Trp Thr Thr Asn Pro Leu Val Glu Tyr Tyr
100 105 110
Ile Val Asp Asn Trp Gly Thr Tyr Arg Pro Pro Gly Gly Gln Gly Tyr
115 120 125
Met Gly Thr Val Thr Ser Asp Gly Ala Thr Tyr Asp Val Tyr Arg Thr
130 135 140
Gln Arg Val Asp Ala Pro Ser Ile Ile Gly Asp His Gln Thr Phe Tyr
145 150 155 160
Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg Thr Gly Gly Thr Ile Thr
165 170 175
Thr Gly Asn His Phe Asp Gly Trp Ala Ser Tyr Gly Met Asn Leu Gly
180 185 190
Thr His Asn Tyr Gln Ile Leu Ala Thr Glu Gly Tyr Gln Ser Ser Gly
195 200 205
Ser Ser Asp Leu Thr Val Ser Glu Gly Ser Ser Ser Ser Ser Gly
210 215 220
Gly Gly Ser Ser Ser Ser Ser Ser Gly Gly Gly Gly Thr Lys Ser Phe
225 230 235 240
Thr Val Arg Ala Arg Gly Thr Ala Gly Gly Glu Ser Ile Thr Leu Arg

Val Asn Asn Gln 245 Val Gln Thr Trp 250 Leu Gly Thr Ser 255 Met Thr
 Asn Tyr Thr Ala 260 Ser Thr Ser Leu Ser Gly Gly Ile Thr Val Ala Phe
 Thr Asn Asp Gly Gly Asn Arg Asp Val Gln Val Asp Tyr Ile Ile Val
 Asn Gly Gln Thr Arg Gln Ser Glu Ala Gln Ser Tyr Asn Thr Gly Leu
 305 310 315 320
 Tyr Ala Asn Gly Arg Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His
 325 335
 Cys Asn Gly Ala Ile Gly Tyr Gly Asn Thr Pro
 340 345

<210> 197
 <211> 636
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 197
 atgtttaagt tcagtaagaa aatgatgacg gttattcttg cagctaccat gagttttggt 60
 ttatttgcaa caacctcaag tgcagcaacc gactattggc aaaattggac cgatggcggc 120
 ggaacggtta atgctgtaaa cggctccggc ggtaattaca gcgtgacatg gcaaaatacc 180
 ggaaattttg tcgtcggcaa aggctggaat accggatcgc ctaaccgaac cattaactac 240
 aatgccggcg tctggcgccc ttccggcaat gggattttga ctctctacgg atggacgaga 300
 aacgcactca ttgaatatta cgctcgtggat agctggggta cttatcggcc tacaggaaca 360
 tataaagggg cggtgacaag tgatgggggc acatatgata tctatacgac catgcggcac 420
 aacgcgcctt ccattgacgg aactcaaacg ttggcccagt actggagtgt tcgacaatcg 480
 aaaagagcga ccgggggtcaa ctctccatt acgttcagca accacgtgaa cgcattgggct 540
 agcaagggaa tgaatctggg aagcagctgg tcatatcagg tgtagctac agagggttat 600
 caaagtagcg gaagctctaa cgtaacagtg tggtaa 636

<210> 198
 <211> 211
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 198
 Met Phe Lys Phe Ser Lys Lys Met Met Thr Val Ile Leu Ala Ala Thr
 1 5 10 15
 Met Ser Phe Gly Leu Phe Ala Thr Thr Ser Ser Ala Ala Thr Asp Tyr
 20 25 30
 Trp Gln Asn Trp Thr Asp Gly Gly Thr Val Asn Ala Val Asn Gly
 35 40 45
 Ser Gly Gly Asn Tyr Ser Val Thr Trp Gln Asn Thr Gly Asn Phe Val
 50 55 60
 Val Gly Lys Gly Trp Asn Thr Gly Ser Pro Asn Arg Thr Ile Asn Tyr
 65 70 75 80
 Asn Ala Gly Val Trp Ala Pro Ser Gly Asn Gly Tyr Leu Thr Leu Tyr
 85 90 95
 Gly Trp Thr Arg Asn Ala Leu Ile Glu Tyr Tyr Val Val Asp Ser Trp
 100 105 110
 Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys Gly Thr Val Thr Ser Asp
 115 120 125
 Gly Gly Thr Tyr Asp Ile Tyr Thr Thr Met Arg His Asn Ala Pro Ser
 130 135 140
 Ile Asp Gly Thr Gln Thr Phe Ala Gln Tyr Trp Ser Val Arg Gln Ser
 145 150 155 160
 Lys Arg Ala Thr Gly Val Asn Ser Ser Ile Thr Phe Ser Asn His Val
 165 170 175

Asn Ala Trp Ala Ser Lys Gly Met Asn Leu Gly Ser Ser Trp Ser Tyr
 180 185 190
 Gln Val Leu Ala Thr Glu Gly Tyr Gln Ser Ser Gly Ser Ser Asn Val
 195 200 205
 Thr Val Trp
 210

<210> 199
 <211> 1074
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 199
 atgatttttc gtctaaagtc gatcacgggc aggcgcgcgc tcgcggcgct ggcctgcctt 60
 gccggcctct acatggcgcc ggcgaaatgc caaacctgca tcacgtcgag ccagacgggc 120
 accaacaacg gcaactactt ttcgttcttg aaagacagcc cgggcacggg gaacttctgc 180
 atgtactccg gcggccgcta cacgtccaac tggagcggca tcaacaactg ggtgggcggc 240
 aagggctggc agacgggctc gtcccgcacc gtctcctact ccggcagctt caattcgccg 300
 ggtaacggct acctgacgct ctacggctgg accaccaatc cgctcatcga gtactacatc 360
 gtcgacaact ggggcagcta tcgtccgcgc ggtggccagg gcttcatggg cacgggtgaac 420
 accgacggcg gcacgtacga catctatcgc acgcaacggg tcaaccagcc gtcgatcatc 480
 ggcaccgcga cgttctacca gtactggagc gtgcggcagt cgaagcgac ccggcggcacc 540
 atcaccacgg ccaaccactt caatgcctgg gccagcctcg gcatgaacct gggacagcac 600
 aactaccagg tgatggccac cgagggctac cagagcagcg gcagctccga catcacgggtg 660
 tgggaaggca cgagcagcgg cgggaagcagc aatggcggca gcagcaacgg ccggcagcagc 720
 aatggtggca gcggcgccac gaagagcttc acggtgcgcg cgcgcggcac tgcgggcggc 780
 gagtccatca cgctgcgggt caacaaccag aacgtgcaga cctggacgct ggggtaccagc 840
 atgcagaact acacggcctc gacctcgctg agcggcggca tcacggtggc gttcaccaac 900
 gacggcggca gccgcgacgt gcaggtggac tacatcatcg tgaatggcca gaccgcagc 960
 tccgaacagc agagctacaa cactggcctc tacgccaatg gaagctgtgg tggcgggttcg 1020
 aacagcgagt ggatgcattg caacggcgcc atcggctacg gcaatacgcc ctga 1074

<210> 200
 <211> 354
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(30)

<400> 200
 Met Ile Phe Gly Leu Lys Ser Ile Thr Gly Arg Arg Ala Val Ala Ala
 1 5 10 15
 Leu Ala Cys Leu Ala Gly Leu Tyr Met Ala Pro Ala Asn Ala Gln Thr
 20 25 30
 Cys Ile Thr Ser Ser Gln Thr Gly Thr Asn Asn Gly Asn Tyr Phe Ser
 35 40 45
 Phe Trp Lys Asp Ser Pro Gly Thr Val Asn Phe Cys Met Tyr Ser Gly
 50 55 60
 Gly Arg Tyr Thr Ser Asn Trp Ser Gly Ile Asn Asn Trp Val Gly Gly
 65 70 75 80
 Lys Gly Trp Gln Thr Gly Ser Ser Arg Thr Val Ser Tyr Ser Gly Ser
 85 90 95
 Phe Asn Ser Pro Gly Asn Gly Tyr Leu Thr Leu Tyr Gly Trp Thr Thr
 100 105 110
 Asn Pro Leu Ile Glu Tyr Tyr Ile Val Asp Asn Trp Gly Ser Tyr Arg
 115 120 125
 Pro Pro Gly Gly Gln Gly Phe Met Gly Thr Val Asn Thr Asp Gly Gly
 130 135 140
 Thr Tyr Asp Ile Tyr Arg Thr Gln Arg Val Asn Gln Pro Ser Ile Ile
 145 150 155 160
 Gly Thr Ala Thr Phe Tyr Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg
 165 170 175

Thr Gly Gly Thr Ile Thr Thr Ala Asn His Phe Asn Ala Trp Ala Ser
 180 185 190
 Leu Gly Met Asn Leu Gly Gln His Asn Tyr Gln Val Met Ala Thr Glu
 195 200 205
 Gly Tyr Gln Ser Ser Gly Ser Ser Asp Ile Thr Val Trp Glu Gly Thr
 210 215 220
 Ser Ser Gly Gly Ser Ser Asn Gly Gly Ser Ser Asn Gly Gly Ser Ser
 225 230 235 240
 Asn Gly Gly Ser Gly Gly Thr Lys Ser Phe Thr Val Arg Ala Arg Gly
 245 250 255
 Thr Ala Gly Gly Glu Ser Ile Thr Leu Arg Val Asn Asn Gln Asn Val
 260 265 270
 Gln Thr Trp Thr Leu Gly Thr Ser Met Gln Asn Tyr Thr Ala Ser Thr
 275 280 285
 Ser Leu Ser Gly Gly Ile Thr Val Ala Phe Thr Asn Asp Gly Gly Ser
 290 295 300
 Arg Asp Val Gln Val Asp Tyr Ile Ile Val Asn Gly Gln Thr Arg Gln
 305 310 315 320
 Ser Glu Gln Gln Ser Tyr Asn Thr Gly Leu Tyr Ala Asn Gly Ser Cys
 325 330 335
 Gly Gly Gly Ser Asn Ser Glu Trp Met His Cys Asn Gly Ala Ile Gly
 340 345 350
 Tyr Gly

<210> 201
 <211> 1002
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 201
 atgaagatga acagctccct cccctccctc cgcgatgtat tcgcgaatga tttccgcattc 60
 ggggcggcgg tcaatcctgt gacgatcgag atgcaaaaac agttgttgat cgatcatgtc 120
 aacagtatta cggcagagaa ccatatgaag tttgagcatc ttcagccgga agaagggaaa 180
 tttacctttc aggaagcggg tccgattgtg gattttgctt gttcgcaccg aatggcgggt 240
 cgagggcaca cacttgatg gcacaaccag actccggatt ggggtgttca agatgggtcaa 300
 ggccatttcg tcagtcggga tgtgttgctt gagcggatga aatgtcacat ttcaactgtt 360
 gtacggcgat acaaggggaaa aatatattgt tgggatgtca tcaacgaagc ggtagccgac 420
 gaaggagacg aattgttgag gccgtcgaag tggcgacaaa tcatcgggga cgattttatg 480
 gaacaagcat ttctctacgc ttatgaagct gacccagatg cactgctttt ttacaatgac 540
 tataatgaat gttttccgga aaagagagaa aaaatttttg cacttgtaaa atcgtgctg 600
 gataaaggca ttccgattca tggcatcggc atgcaggcgc actggagcct gaccgcgccg 660
 tcgcttgatg aaattcgtgc ggcgattgaa cggatgctg cccttggtgt tgttcttcat 720
 attacggaac tcgatgtatc catgtttgaa ttacacgatc gtcgaaccga tttggctgtc 780
 ccgacgaacg aaatgatcga acagcaagca gaacgggatg ggcaaatatt tgctttgttt 840
 aaggagatga gcgatgttat tcaaagtgtc acattttggg gaattgctga tgaccataca 900
 tggctcgata actttccagt gcacgggaga aaaaactggc cgcttttggt cgatgaacag 960
 cataaaccga aaccagcttt ttggcgggca gtgagtgtct ga 1002

<210> 202
 <211> 333
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 202
 Met Lys Met Asn Ser Ser Leu Pro Ser Leu Arg Asp Val Phe Ala Asn
 1 5 10 15
 Asp Phe Arg Ile Gly Ala Ala Val Asn Pro Val Thr Ile Glu Met Gln
 20 25 30
 Lys Gln Leu Ile Asp His Val Asn Ser Ile Thr Ala Glu Asn His
 35 40 45
 Met Lys Phe Glu His Leu Gln Pro Glu Glu Gly Lys Phe Thr Phe Gln
 50 55 60

Glu Ala Asp Arg Ile Val Asp Phe Ala Cys Ser His Arg Met Ala Val
 65 70 75 80
 Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro Asp Trp Val Phe
 85 90 95
 Gln Asp Gly Gln Gly His Phe Val Ser Arg Asp Val Leu Leu Glu Arg
 100 105 110
 Met Lys Cys His Ile Ser Thr Val Arg Arg Tyr Lys Gly Lys Ile
 115 120 125
 Tyr Cys Trp Asp Val Ile Asn Glu Ala Val Ala Asp Glu Gly Asp Glu
 130 135 140
 Leu Leu Arg Pro Ser Lys Trp Arg Gln Ile Ile Gly Asp Asp Phe Met
 145 150 155 160
 Glu Gln Ala Phe Leu Tyr Ala Tyr Glu Ala Asp Pro Asp Ala Leu Leu
 165 170 175
 Phe Tyr Asn Asp Tyr Asn Glu Cys Phe Pro Glu Lys Arg Glu Lys Ile
 180 185 190
 Phe Ala Leu Val Lys Ser Leu Arg Asp Lys Gly Ile Pro Ile His Gly
 195 200 205
 Ile Gly Met Gln Ala His Trp Ser Leu Thr Arg Pro Ser Leu Asp Glu
 210 215 220
 Ile Arg Ala Ala Ile Glu Arg Tyr Ala Ser Leu Gly Val Val Leu His
 225 230 235 240
 Ile Thr Glu Leu Asp Val Ser Met Phe Glu Phe His Asp Arg Arg Thr
 245 250 255
 Asp Leu Ala Val Pro Thr Asn Glu Met Ile Glu Gln Gln Ala Glu Arg
 260 265 270
 Tyr Gly Gln Ile Phe Ala Leu Phe Lys Glu Tyr Arg Asp Val Ile Gln
 275 280 285
 Ser Val Thr Phe Trp Gly Ile Ala Asp Asp His Thr Trp Leu Asp Asn
 290 295 300
 Phe Pro Val His Gly Arg Lys Asn Trp Pro Leu Leu Phe Asp Glu Gln
 305 310 315 320
 His Lys Pro Lys Pro Ala Phe Trp Arg Ala Val Ser Val
 325 330

<210> 203
 <211> 687
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 203
 atgaaatctg caccgcgact tttggtggcg ctatcacgca tacttccgat cgcacttgtg 60
 ctgttgctcg ccccgctccc cgcgcaagcc caacaggtct gcaacaacgg aacgggcacg 120
 cataacggct tcttctggac gttttggaag gacggcggca cggcctgcat gacgctcggc 180
 tcgggcggca attatagcac gacgttcaat ctgtccggcg gccgcaacct tgttcggggc 240
 aagggtggc agactggctc caccaaccga gtcgtcgggt acaatgcggg cgtctggaac 300
 ccaggcacca attcttatct gacgtcttat ggctggtcga cgaatccgct cgtcgaatat 360
 tatgtcgtgg accattgggg cagccaattc accccgccag gcaacggcgc gcagagcatg 420
 gggaccgtga ccaccgacgg cggcacctac aacatctacc gcacccaacg cgtcaacgcg 480
 ctttcgatca tcggcaacgc cagtttctac caatattgga gcgtgcgcac ttcgcgccgc 540
 gggcaaggca cgaacaacac gatcaccttc gccaatcacg tcaacgcttg gcgcagccgc 600
 ggcataaacc ttgggaccat gaattatcaa gtcattggcca cggaagggtt cggctcgaac 660
 ggaagctcca acctcacagt atggttag 687

<210> 204
 <211> 228
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(30)

<400> 204

Met Lys Ser Ala Arg Ala Leu Leu Val Ala Leu Ser Arg Ile Leu Pro
 1 5 10 15
 Ile Ala Leu Val Leu Leu Ala Pro Val Pro Ala Gln Ala Gln Gln
 20 25 30
 Val Cys Asn Asn Gly Thr Gly Thr His Asn Gly Phe Phe Trp Thr Phe
 35 40 45
 Trp Lys Asp Gly Gly Thr Ala Cys Met Thr Leu Gly Ser Gly Gly Asn
 50 55 60
 Tyr Ser Thr Thr Phe Asn Leu Ser Gly Gly Arg Asn Leu Val Ala Gly
 65 70 75 80
 Lys Gly Trp Gln Thr Gly Ser Thr Asn Arg Val Val Gly Tyr Asn Ala
 85 90 95
 Gly Val Trp Asn Pro Gly Thr Asn Ser Tyr Leu Thr Leu Tyr Gly Trp
 100 105 110
 Ser Thr Asn Pro Leu Val Glu Tyr Tyr Val Val Asp His Trp Gly Ser
 115 120 125
 Gln Phe Thr Pro Pro Gly Asn Gly Ala Gln Ser Met Gly Thr Val Thr
 130 135 140
 Thr Asp Gly Gly Thr Tyr Asn Ile Tyr Arg Thr Gln Arg Val Asn Ala
 145 150 155 160
 Pro Ser Ile Ile Gly Asn Ala Thr Phe Tyr Gln Tyr Trp Ser Val Arg
 165 170 175
 Thr Ser Arg Arg Gly Gln Gly Thr Asn Asn Thr Ile Thr Phe Ala Asn
 180 185 190
 His Val Asn Ala Trp Arg Ser Arg Gly Met Asn Leu Gly Thr Met Asn
 195 200 205
 Tyr Gln Val Met Ala Thr Glu Phe Gly Ser Asn Gly Ser Ser Asn
 210 215 220
 Leu Thr Val Trp
 225

<210> 205
 <211> 1068
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 205
 atgcaaattt tcaaatacacc actgtcatgg gccggatcac tattactgat cctgtccacc 60
 gccctgtttt caacagcggc cactgcccag gaatactgct ccaaccagac cggtacacac 120
 agcggttttt actttaccct ttggtctgac ggcggcggta ctgcctgcat tactctggga 180
 gacgacggaa attacagtta cacctgggtcc aacacaggca attttgtcgg tggcaagggc 240
 tggagtaccg gcacctccaa tcgggtgatc ggttacaacg ccggagacta ctgccctcc 300
 ggcaactcct acctggcgct gtatggctgg agcaccaatc cactgattga gtactacgtg 360
 gtggatagct ggggtagctg gcgtccgccg ggtggcacct cggtaggtac agtcaccagc 420
 gatggcggga ctacgacct gtaccgcacc gagcgcgtgc agcagccctc catcgaaggc 480
 acggccacct tctatcaata ttggagcgtg cgcacctcac agcgtcccca ggggcagaac 540
 aacaccatca cctttcagaa ccacgtggat gcctgggcca atcagggctg gaacctcggc 600
 acccacaact atcaggtaat ggcgaccgaa ggctacgaaa gcagcggcag ctccaacgtc 660
 acggtttggg attccggcac cagtagcggg aacgggtggc acgctggcgg cggtgggtggc 720
 gaggcaggta acggctccaa ctactgggtc gtgcgtgcgg tgggcacttc gggcaacgaa 780
 cagttgcgcg tcaacgtcag cggcaacacg gttgaaaccc tgaacctgtc taccaactgg 840
 caggactaca ccatcaacac caacgcttcc ggcatgtga atgtggagtt gatcaacgat 900
 cagggcgagg gctacgaagc ccgggtggaa tacgtcatcg tcaacggcga taccgcgtac 960
 ggcgctgatc agagctacaa caccagcgcc tgggacggcg agtgcggcgg cggttccttt 1020
 accatgtgga tgcactgcga aggcattctc ggttttggcg atatgtaa 1068

<210> 206
 <211> 355
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(29)

<400> 206
Met Gln Ile Phe Lys Ser Pro Leu Ser Trp Ala Gly Ser Leu Leu Leu
1 5 10 15
Ile Leu Ser Thr Ala Leu Phe Ser Thr Ala Ala Thr Ala Gln Glu Tyr
20 25 30
Cys Ser Asn Gln Thr Gly Thr His Ser Gly Phe Tyr Phe Thr His Trp
35 40 45
Ser Asp Gly Gly Gly Thr Ala Cys Ile Thr Leu Gly Asp Asp Gly Asn
50 55 60
Tyr Ser Tyr Thr Trp Ser Asn Thr Gly Asn Phe Val Gly Gly Lys Gly
65 70 75 80
Trp Ser Thr Gly Thr Ser Asn Arg Val Ile Gly Tyr Asn Ala Gly Asp
85 90 95
Tyr Ser Pro Ser Gly Asn Ser Tyr Leu Ala Leu Tyr Gly Trp Ser Thr
100 105 110
Asn Pro Leu Ile Glu Tyr Tyr Val Asp Ser Trp Gly Ser Trp Arg
115 120 125
Pro Pro Gly Gly Thr Ser Val Gly Thr Val Thr Ser Asp Gly Gly Thr
130 135 140
Tyr Asp Leu Tyr Arg Thr Glu Arg Val Gln Gln Pro Ser Ile Glu Gly
145 150 155 160
Thr Ala Thr Phe Tyr Gln Tyr Trp Ser Val Arg Thr Ser Gln Arg Pro
165 170 175
Gln Gly Gln Asn Asn Thr Ile Thr Phe Gln Asn His Val Asp Ala Trp
180 185 190
Ala Asn Gln Gly Trp Asn Leu Gly Thr His Asn Tyr Gln Val Met Ala
195 200 205
Thr Glu Gly Tyr Glu Ser Ser Gly Ser Ser Asn Val Thr Val Trp Asp
210 215 220
Ser Gly Thr Ser Ser Gly Asn Gly Gly Asn Ala Gly Gly Gly Gly Gly
225 230 235 240
Glu Ala Gly Asn Gly Ser Asn Ser Leu Val Val Arg Ala Val Gly Thr
245 250 255
Ser Gly Asn Glu Gln Leu Arg Val Asn Val Ser Gly Asn Thr Val Glu
260 265 270
Thr Leu Asn Leu Ser Thr Asn Trp Gln Asp Tyr Thr Ile Asn Thr Asn
275 280 285
Ala Ser Gly Asp Val Asn Val Glu Leu Ile Asn Asp Gln Gly Glu Gly
290 295 300
Tyr Glu Ala Arg Val Glu Tyr Val Ile Val Asn Gly Asp Thr Arg Tyr
305 310 315 320
Gly Ala Asp Gln Ser Tyr Asn Thr Ser Ala Trp Asp Gly Glu Cys Gly
325 330 335
Gly Gly Ser Phe Thr Met Trp Met His Cys Glu Gly Ile Leu Gly Phe
340 345 350
Gly Asp Met
355

<210> 207
<211> 633
<212> DNA
<213> Unknown

<220>
<223> obtained from an environmental sample

<400> 207
atgaaattaa aaaagaagat gctcacttta ctctgacgg cttcgatgag tttcggttta 60
tttggggcaa cctcgagtgc agcaacggat tattggcaat attggacgga tggcggcgga 120
acggtgaatg cggttaacgg gtccgggggc aattacagcg taacttggca aaatagcggg 180
aacttcgttg tcggcaaaag ctggagcgta gggtcgcca atcggacgat caattacaat 240
gccggcatct gggaaccttc ggggaacggg tacttgacct tttacggaatg gactagaaac 300
tcgctgatcg agtattacgt tgtcgacagt tgggggacgt accggccaac aggtactcac 360
aaaggaacgg tgaacagcga cggaggcacc tacgatattt atacgaccat gcgctataat 420
gcgcttcca ttgatggcac gcagacgttc caacagttct ggagcgtgcg gcaatcgaaa 480
cgaccaaccg gcagcaacgt ctccatcacc ttacagcaatc acgtgaatgc ctggagaagc 540
aagggcatag acctgggcag cagctgggtcg taccaggctc tggcgacgga aggctatcac 600
agcagcggaa gatccaacgt cacgggtgtgg taa 633

<210> 208
 <211> 210
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(27)

<400> 208
 Met Lys Leu Lys Lys Lys Met Leu Thr Leu Leu Leu Thr Ala Ser Met
 1 5 10 15
 Ser Phe Gly Leu Phe Gly Ala Thr Ser Ser Ala Ala Thr Asp Tyr Trp
 20 25 30
 Gln Tyr Trp Thr Asp Gly Gly Gly Thr Val Asn Ala Val Asn Gly Ser
 35 40 45
 Gly Gly Asn Tyr Ser Val Thr Trp Gln Asn Ser Gly Asn Phe Val Val
 50 55 60
 Gly Lys Gly Trp Ser Val Gly Ser Pro Asn Arg Thr Ile Asn Tyr Asn
 65 70 75 80
 Ala Gly Ile Trp Glu Pro Ser Gly Asn Gly Tyr Leu Thr Leu Tyr Gly
 85 90 95
 Trp Thr Arg Asn Ser Leu Ile Glu Tyr Tyr Val Val Asp Ser Trp Gly
 100 105 110
 Thr Tyr Arg Pro Thr Gly Thr His Lys Gly Thr Val Asn Ser Asp Gly
 115 120 125
 Gly Thr Tyr Asp Ile Tyr Thr Met Arg Tyr Asn Ala Pro Ser Ile
 130 135 140
 Asp Gly Thr Gln Thr Phe Gln Gln Phe Trp Ser Val Arg Gln Ser Lys
 145 150 155 160
 Arg Pro Thr Gly Ser Asn Val Ser Ile Thr Phe Ser Asn His Val Asn
 165 170 175
 Ala Trp Arg Ser Lys Gly Met Asn Leu Gly Ser Ser Trp Ser Tyr Gln
 180 185 190
 Val Leu Ala Thr Glu Gly Tyr Gln Ser Ser Gly Arg Ser Asn Val Thr
 195 200 205
 Val Trp
 210

<210> 209
 <211> 1194
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 209
 atgaaaacat ttagtgtagc caagtctagc gttgttttcg caatggcttt gggatatggct 60
 tcgacagctt ttgctcagga tttctgcagc aatgcgcaac attccggcca aaaggtaacg 120
 attacttcga accaaactgg taaaatcggc gatatacggtt acgaactctg ggacgaaaac 180
 ggtcatgggtg gtagtgctac cttctatagc gatgggttcca tggactgcaa tatcactgggt 240
 gctaaggact atctctgccc tgcgggcctt tccctcggca gtaacaagac ctacaaggaa 300
 cttgggtgggtg atatgattgc cgagttcaag cttgtgaaga gcggtgcccc gaatgtgggt 360
 tactcttata tcgggtatcta tggctggatg gaaggtgttt ctggaacgcc tagccagttg 420
 gtcgaataact acgtgattga taacaccctc gccaatgaca tgccgggtag ctggattggt 480
 aacgaaagaa agggtaaccat tacggttgac ggcggtacct atactgttta tcgcaatacc 540
 cgtacaggtc cggctattaa gaacagcggg aacgtcacgt tctatcagta tttcagcgtt 600
 cgtacctctc cgcgcgattg cgggtaccatc aatatttccg aacacatgag acagtgggaa 660
 aagatgggca tgaccatggg taagctctac gaagccaagg tgcttggcga agcgggtaac 720
 gtgaatggcg aagtccgcgg tggtcacatg gacttcccgc atgctaagggt ttatgtgaaa 780
 aacggctctg atccggcttc ttctcttct gtgaagtcca gctcttctac agtaacgcca 840
 aaatccagct cctcgaaggg taacggcaac gtttctggta aaattgacgc ctgcaaggac 900
 gctatgggcc atgaaggcaa agaacaactc tagcgtgacg tagcgtgacg 960
 ggtaacgtcg gcagctctcc gtaccactat gaaatttggg atcaggggtg taacaactcc 1020
 atgacgttct acgacaacgg tacttataag gcaagctgga atggtaccaa cgacttcctt 1080

gctcgtgtcg gtttcaagta tgatgaaaag cacacttacg aagaacttgg ccctatcgat 1140
gcctactaca agtggagcaa gcagggtagt gctggtggct acaactacat cggc 1194

<210> 210
<211> 398
<212> PRT
<213> Unknown

<220>
<223> obtained from an environmental sample

<221> SIGNAL
<222> (1)...(25)

<400> 210
Met Lys Thr Phe Ser Val Thr Lys Ser Ser Val Val Phe Ala Met Ala
1 5 10 15
Leu Gly Met Ala Ser Thr Ala Phe Ala Gln Asp Phe Cys Ser Asn Ala
20 25 30
Gln His Ser Gly Gln Lys Val Thr Ile Thr Ser Asn Gln Thr Gly Lys
35 40 45
Ile Gly Asp Ile Gly Tyr Glu Leu Trp Asp Glu Asn Gly His Gly Gly
50 55 60
Ser Ala Thr Phe Tyr Ser Asp Gly Ser Met Asp Cys Asn Ile Thr Gly
65 70 75 80
Ala Lys Asp Tyr Leu Cys Arg Ala Gly Leu Ser Leu Gly Ser Asn Lys
85 90 95
Thr Tyr Lys Glu Leu Gly Gly Asp Met Ile Ala Glu Phe Lys Leu Val
100 105 110
Lys Ser Gly Ala Gln Asn Val Gly Tyr Ser Tyr Ile Gly Ile Tyr Gly
115 120 125
Trp Met Glu Gly Val Ser Gly Thr Pro Ser Gln Leu Val Glu Tyr Tyr
130 135 140
Val Ile Asp Asn Thr Leu Ala Asn Asp Met Pro Gly Ser Trp Ile Gly
145 150 155 160
Asn Glu Arg Lys Gly Thr Ile Thr Val Asp Gly Gly Thr Tyr Thr Val
165 170 175
Tyr Arg Asn Thr Arg Thr Gly Pro Ala Ile Lys Asn Ser Gly Asn Val
180 185 190
Thr Phe Tyr Gln Tyr Phe Ser Val Arg Thr Ser Pro Arg Asp Cys Gly
195 200 205
Thr Ile Asn Ile Ser Glu His Met Arg Gln Trp Glu Lys Met Gly Met
210 215 220
Thr Met Gly Lys Leu Tyr Glu Ala Lys Val Leu Gly Glu Ala Gly Asn
225 230 235 240
Val Asn Gly Glu Val Arg Gly Gly His Met Asp Phe Pro His Ala Lys
245 250 255
Val Tyr Val Lys Asn Gly Ser Asp Pro Ala Ser Ser Ser Ser Val Lys
260 265 270
Ser Ser Ser Ser Thr Val Thr Pro Lys Ser Ser Ser Ser Lys Gly Asn
275 280 285
Gly Asn Val Ser Gly Lys Ile Asp Ala Cys Lys Asp Ala Met Gly His
290 295 300
Glu Gly Lys Glu Thr Arg Thr Gln Gly Gln Asn Asn Ser Ser Val Thr
305 310 315 320
Gly Asn Val Gly Ser Pro Tyr His Tyr Glu Ile Trp Tyr Gln Gly
325 330 335
Gly Asn Asn Ser Met Thr Phe Tyr Asp Asn Gly Thr Tyr Lys Ala Ser
340 345 350
Trp Asn Gly Thr Asn Asp Phe Leu Ala Arg Val Gly Phe Lys Tyr Asp
355 360 365
Glu Lys His Thr Tyr Glu Glu Leu Gly Pro Ile Asp Ala Tyr Tyr Lys
370 375 380
Trp Ser Lys Gln Gly Ser Ala Gly Gly Tyr Asn Tyr Ile Gly
385 390 395

<210> 211
<211> 1086
<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 211

atgataagtt	ctaaagcatc	acagtcattg	ggctgggtcac	tattgggtggc	cctgtccgcc	60
gtttctgctt	cggcgacagc	ttccgcccag	caacactgct	ccaaccaaac	cggtacgcac	120
aacggttttt	actttaccca	ttggtcagac	gggtggcggt	ccgcctgcat	gactctgggg	180
gacgacggca	actacagcta	tacctgggtc	aacactggca	attttgtcgg	tggttaagggc	240
tggagcacag	gtacatccaa	ccgggtgatt	ggttacaacg	ccggagacta	ctcgccctcc	300
ggcaactcct	acctggcact	gtatggctgg	agcaccaatc	cgctgattga	atattacgtg	360
gtcgcacagt	ggggcagctg	gcgtccgccg	gggtggcacct	ctgtgggcac	ggtaaccagc	420
gacggtggca	cttacgacct	gtaccgaacc	cagcgtgtgc	agcagccctc	cattgagggg	480
acggccacct	tctatcaata	ctggagcgtg	cgacactcac	agcggcctca	ggggcaaac	540
aacaccatca	cctttcagaa	ccacgtgaat	gcctggggcca	atcagggctg	gaatctgggc	600
acccacaact	atcaggtgat	ggcgaccgaa	ggctacgaaa	gcagcggcag	ctccaacgtc	660
accgtttggg	attccggcac	cagtagcggg	ggcggtggcg	gtggcaacgc	gggcggcgcc	720
ggagcccccg	gtggtggtga	ggctggaggg	ggctccaact	cactggttgt	gcgtgcgggt	780
ggcacttcgg	gcaatgaaca	gttgccgctc	aacgtcagtg	gcaacacggg	ggaaaccctg	840
aacctgtcta	ccaactggca	ggactacacc	atcaacacca	acgcctccgg	cgatgtcaat	900
gtggaattga	tcaacgacca	gggcgaaggc	tacgagggcc	gcgtcgagta	cgatcatcatc	960
aacggcgata	cccgtacagg	cggcgaccag	agctacaaca	ccagcgctcg	ggacggcgag	1020
tgcggtagcg	gttcctttac	catgtggatg	cactgcgaag	gcacccctcg	ttttggcgat	1080
atgtaa						1086

<210> 212

<211> 361

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<221> SIGNAL

<222> (1)...(29)

<400> 212

Met	Ile	Ser	Ser	Lys	Ala	Ser	Gln	Ser	Trp	Gly	Trp	Ser	Leu	Leu	Val
1				5					10					15	
Ala	Leu	Ser	Ala	Val	Leu	Leu	Ser	Ala	Thr	Ala	Ser	Ala	Gln	Gln	His
			20					25					30		
Cys	Ser	Asn	Gln	Thr	Gly	Thr	His	Asn	Gly	Phe	Tyr	Phe	Thr	His	Trp
		35					40					45			
Ser	Asp	Gly	Gly	Gly	Thr	Ala	Cys	Met	Thr	Leu	Gly	Asp	Asp	Gly	Asn
	50					55					60				
Tyr	Ser	Tyr	Thr	Trp	Ser	Asn	Thr	Gly	Asn	Phe	Val	Gly	Gly	Lys	Gly
	65				70					75				80	
Trp	Ser	Thr	Gly	Thr	Ser	Asn	Arg	Val	Ile	Gly	Tyr	Asn	Ala	Gly	Asp
			85					90						95	
Tyr	Ser	Pro	Ser	Gly	Asn	Ser	Tyr	Leu	Ala	Leu	Tyr	Gly	Trp	Ser	Thr
		100						105					110		
Asn	Pro	Leu	Ile	Glu	Tyr	Tyr	Val	Val	Asp	Ser	Trp	Gly	Ser	Trp	Arg
		115					120					125			
Pro	Pro	Gly	Gly	Thr	Ser	Val	Gly	Thr	Val	Thr	Ser	Asp	Gly	Gly	Thr
	130					135					140				
Tyr	Asp	Leu	Tyr	Arg	Thr	Gln	Arg	Val	Gln	Gln	Pro	Ser	Ile	Glu	Gly
	145				150					155				160	
Thr	Ala	Thr	Phe	Tyr	Gln	Tyr	Trp	Ser	Val	Arg	Thr	Ser	Gln	Arg	Pro
			165					170						175	
Gln	Gly	Gln	Asn	Asn	Thr	Ile	Thr	Phe	Gln	Asn	His	Val	Asn	Ala	Trp
			180					185					190		
Ala	Asn	Gln	Gly	Trp	Asn	Leu	Gly	Thr	His	Asn	Tyr	Gln	Val	Met	Ala
		195					200					205			
Thr	Glu	Gly	Tyr	Glu	Ser	Ser	Gly	Ser	Ser	Asn	Val	Thr	Val	Trp	Asp
	210				215						220				
Ser	Gly	Thr	Ser	Ser	Gly	Gly	Gly	Gly	Gly	Gly	Asn	Ala	Gly	Gly	Gly
	225				230					235				240	
Gly	Ala	Pro	Gly	Gly	Gly	Glu	Ala	Gly	Gly	Gly	Ser	Asn	Ser	Leu	Val

Val Arg Ala Val 245 Gly Thr Ser Gly Asn 250 Glu Gln Leu Arg Val 255 Asn Val
 Ser Gly Asn Thr Val 260 Glu Thr Leu Asn 265 Leu Ser Thr Asn 270 Trp Gln Asp
 Tyr Thr 275 Ile Asn Thr Asn 280 Ser Gly Asp Val 285 Asn Val Glu Leu Ile
 Asn Asp Gln Gly Glu 290 Gly Tyr Glu Ala Arg Val 300 Glu Tyr Val Ile Ile
 305 Asn Gly Asp Thr Arg 310 Tyr Gly Ala Asp Gln Ser Tyr Asn Thr Ser Ala
 Trp Asp Gly Glu 315 Cys Gly Ser Gly Ser 320 Phe Thr Met Trp Met His Cys
 Glu Gly Ile Leu Gly Phe Gly Asp Met 335 350
 355 360

<210> 213
 <211> 912
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 213
 gtgaacgcac aacaaaccct tacgtctaac tccaccggta ctcatggtgg tcactactat 60
 tctttctgga aggactccgg caatgcgtcc ttcactctct acgatggcgg acgttacggc 120
 tcgcaatgga atagcggcac caacaattgg gtgggcggtta aaggctggaa cccgggcggc 180
 gcaaaagtcg ttaactacga aggttattac ggcgttaaca attcccagaa ttcttacctg 240
 gcactctacg ggtggaccgg caatccgctg atcgagtact acataatcga aagttacggg 300
 tcgtacaacc catcgagctg tagtggcggt actaactacg gtagcttcca aagcgatggg 360
 gcgacctata acgtccgccc ttgccagcgc gtacagcagc catcgattga tggaacgcaa 420
 acgtttctatc agtatttcag cgttcgctca cccaaaaagg gcttcggcca aatcagcggc 480
 actatcaatg taggcaacca cttaattat tggggccagca aagggtgaa tttgggtagc 540
 cagattaca tggttctggc gactgaaggc tatcagagca gcggcaattc agatatttcc 600
 gtgtccgaag gcagcagcgg cggctcttcc tcaggcggtt cgacctccag cggaagctcc 660
 tccggtagta cgaccagttc ttcaggaggc ggtggcgggc gcatcacagt acgtgctcgc 720
 ggcaactaat gtgatgagcg tatcagcctg cgtgtcgggc gttctgcggg agccagttgg 780
 acactcagta ccagcgacac aagctatagc tacacaggcg gcgctcttgg cgatatccag 840
 gtggaattcg atatcaagct tatcgatacc gtcgacctcg agggggggcc cggatcccaa 900
 ttcgccctat ag 912

<210> 214
 <211> 303
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 214
 Val Asn Ala Gln Gln Thr Leu Thr Ser Asn Ser Thr Gly Thr His Gly
 1 5 10 15
 Gly His Tyr Tyr Ser Phe Trp Lys Asp Ser Gly Asn Ala Ser Phe Thr
 20 25 30
 Leu Tyr Asp Gly Gly Arg Tyr Gly Ser Gln Trp Asn Ser Gly Thr Asn
 35 40 45
 Asn Trp Val Gly Gly Lys Gly Trp Asn Pro Gly Gly Ala Lys Val Val
 50 55 60
 Asn Tyr Glu Gly Tyr Tyr Gly Val Asn Asn Ser Gln Asn Ser Tyr Leu
 65 70 75 80
 Ala Leu Tyr Gly Trp Thr Arg Asn Pro Leu Ile Glu Tyr Tyr Ile Ile
 85 90 95
 Glu Ser Tyr Gly Ser Tyr Asn Pro Ser Ser Cys Ser Gly Gly Thr Asn
 100 105 110
 Tyr Gly Ser Phe Gln Ser Asp Gly Ala Thr Tyr Asn Val Arg Arg Cys
 115 120 125
 Gln Arg Val Gln Gln Pro Ser Ile Asp Gly Thr Gln Thr Phe Tyr Gln
 130 135 140

Tyr Phe Ser Val Arg Ser Pro Lys Lys Gly Phe Gly Gln Ile Ser Gly
 145 150 155 160
 Thr Ile Asn Val Gly Asn His Phe Asn Tyr Trp Ala Ser Lys Gly Leu
 165 170 175
 Asn Leu Gly Ser His Asp Tyr Met Val Leu Ala Thr Glu Gly Tyr Gln
 180 185 190
 Ser Ser Gly Asn Ser Asp Ile Ser Val Ser Glu Gly Ser Ser Gly Gly
 195 200 205
 Ser Ser Ser Gly Gly Ser Thr Ser Ser Gly Ser Ser Ser Gly Ser Thr
 210 215 220
 Thr Ser Ser Ser Gly Gly Gly Gly Gly Ile Thr Val Arg Ala Arg
 225 230 235 240
 Gly Thr Asn Gly Asp Glu Arg Ile Ser Leu Arg Val Gly Gly Ser Ala
 245 250 255
 Val Ala Ser Trp Thr Leu Ser Thr Ser Ala Gln Ser Tyr Ser Tyr Thr
 260 265 270
 Gly Gly Ala Ser Gly Asp Ile Gln Val Glu Phe Asp Ile Lys Leu Ile
 275 280 285
 Asp Thr Val Asp Leu Glu Gly Gly Pro Gly Thr Gln Phe Ala Leu
 290 295 300

<210> 215
 <211> 1065
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 215
 atgtttgcaa gattcgagaa actggccgcg gcgggtaaa cgtcgtggc cctggcagg 60
 ctgcgccctt tgggcacggc gcctgccaat gcacagacct gtctcacgaa caattccacc 120
 ggcaccaaca acggctacta ctactcgttc tggaggaca gcggcaacgt gaccttctgc 180
 atgtacgggg gcggccgcta tacctcgagc tggagcaaca tcaacaactg ggtgggcggc 240
 aagggctgga atccgggcgg tcgtcggacc gtcacctatt cggggacgtt caaccgaac 300
 ggcaattcct atctcacgct gtacggctgg accaccaatc cactggtcga gtactacatc 360
 gtcgacagct ggggcagctg gcgtccgcgc ggttccggct acatgggttc cgtcacgagc 420
 gacggcggca cctacgacat ctatcgacag cagcgcgtca accagccctc gatcatcggc 480
 accgcgacgt tctaccagta ctggagcgtg cggcagcaga agcgcgtggg tggcaccatc 540
 accaccggca accacttcca tgcctgggct tcgctgggca tgaacctcgg ccagcacaac 600
 tacatggtca tggccaccga gggctaccag agcagcggca gctccgacat cacggtgggc 660
 ggcaccagca gctcctcgtc gtcgagcggg ggcagcagca gcagtagcag cagcagcggg 720
 ggtggcggct cgaagagctt caccgtgcgc gcgcgggggt cgacgggcgg tgagcagatc 780
 agtttgcgcg tgaacaacca gaccgtgcag aactggacgc tgggcaccag catgcagaac 840
 tacaccgcgt ccaccaacct gagcggcggc atcacctgac acttcaccaa tgacagcggc 900
 aaccgcgacg tgcaggtgga ctacatccag gtgaacggcc agacgcgtca atccgagcag 960
 cagagctaca acaccgggct gtatgccaac ggcagctgtg gcggcggcgg ctacagcgag 1020
 tggatgcatt gcaatggcgc gatcggttac ggcaacacgc cgtag 1065

<210> 216
 <211> 354
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(31)

<400> 216
 Met Phe Ala Arg Phe Glu Lys Leu Ala Ala Ala Gly Lys Ala Val Val
 1 5 10 15
 Ala Leu Ala Gly Leu Ala Leu Leu Gly Thr Ala Pro Ala Asn Ala Gln
 20 25 30
 Thr Cys Leu Thr Asn Asn Ser Thr Gly Thr Asn Asn Gly Tyr Tyr Tyr
 35 40 45
 Ser Phe Trp Lys Asp Ser Gly Asn Val Thr Phe Cys Met Tyr Gly Gly
 50 55 60

Gly Arg Tyr Thr Ser Gln Trp Ser Asn Ile Asn Asn Trp Val Gly Gly
 65 70 75 80
 Lys Gly Trp Asn Pro Gly Gly Arg Arg Thr Val Thr Tyr Ser Gly Thr
 85 90 95
 Phe Asn Pro Asn Gly Asn Ser Tyr Leu Thr Leu Tyr Gly Trp Thr Thr
 100 105 110
 Asn Pro Leu Val Glu Tyr Tyr Ile Val Asp Ser Trp Gly Ser Trp Arg
 115 120 125
 Pro Pro Gly Ser Gly Tyr Met Gly Ser Val Thr Ser Asp Gly Gly Thr
 130 135 140
 Tyr Asp Ile Tyr Arg Thr Gln Arg Val Asn Gln Pro Ser Ile Ile Gly
 145 150 155 160
 Thr Ala Thr Phe Tyr Gln Tyr Trp Ser Val Arg Gln Gln Lys Arg Val
 165 170 175
 Gly Gly Thr Ile Thr Thr Gly Asn His Phe Asp Ala Trp Ala Ser Leu
 180 185 190
 Gly Met Asn Leu Gly Gln His Asn Tyr Met Val Met Ala Thr Glu Gly
 195 200 205
 Tyr Gln Ser Ser Gly Ser Ser Asp Ile Thr Val Gly Gly Thr Ser Ser
 210 215 220
 Ser Ser Ser Ser Gly Gly Ser Ser Ser Ser Ser Ser Ser Ser Gly
 225 230 235 240
 Gly Gly Gly Ser Lys Ser Phe Thr Val Arg Ala Arg Gly Ser Thr Gly
 245 250 255
 Gly Glu Gln Ile Ser Leu Arg Val Asn Asn Gln Thr Val Gln Asn Trp
 260 265 270
 Thr Leu Gly Thr Ser Met Gln Asn Tyr Thr Ala Ser Thr Asn Leu Ser
 275 280 285
 Gly Gly Ile Thr Val His Phe Thr Asn Asp Ser Gly Asn Arg Asp Val
 290 295 300
 Gln Val Asp Tyr Ile Gln Val Asn Gly Gln Thr Arg Gln Ser Glu Gln
 305 310 315 320
 Gln Ser Tyr Asn Thr Gly Leu Tyr Ala Asn Gly Ser Cys Gly Gly Gly
 325 330 335
 Gly Tyr Ser Glu Trp Met His Cys Asn Gly Ala Ile Gly Tyr Gly Asn
 340 345 350
 Thr Pro

<210> 217
 <211> 1083
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 217
 atgactttcg tcaagacgat caccggcaga cgcgccatcg cggcggttct ctgcctcgcc 60
 ggcctctaca tggcgccggc aaacgcgcaa acctgcatca cgtccagcca gaccggcacc 120
 aacaacggga actacttttc gttctgga aaacgcgcaa gacagcccgg gcacgggtgaa cttctgcatg 180
 taccggaatg gccgctacac ctggaactgg agcgccatca acaactgggt cggcggcaag 240
 ggctgggtcga ccggctccag ccgcaccgtc agctattcgg gcagcttcaa ttcgcccggc 300
 aacggctacc tgaactctta cgggtggacc accaaccgc tcatcgagta ctacatcgtc 360
 gagaactggg gtaactaccg cccgccgggc ggccaggggt acatggggac cgtcaattcc 420
 gacggggcga cctatgacat ctaccggacc ttccgggaca accagccctg catcacgggc 480
 aactcctgag acttctacca gtactggagc gtgcgccagt ccaagcgag cagcggcacc 540
 atcaccacgg ccaatcacct cgcgccgtgg aacagcctcg gcatgaacct gggccagcac 600
 aactaccagg tcatggccac cgaggggtac cagagcagcg gcagctccga catcacggtc 660
 acggaaggcg gcggcgccag cagcaatggg ggcagcagca acggcgccag cagcaatggc 720
 ggcagcagca atggcgccg cggcgccacc aagagcttca cgggtccgag ccgtggcacc 780
 gcgggtggcg agtccatcac gctgcgtgac aacaaccaga acgtgcagac ctggacgctg 840
 ggcaccggca tgcagaacta cacggcctcg acctcgctga gcgggtggcat cacggtgcac 900
 ttcaccaacg acggcggaag ccgcgacgtg caggtggact acatccaggt gaacggcagc 960
 acgcgccagt ccgaggcaca gagctacaac accggcgcc ctctgaacgg ccgttgcggc 1020
 ggtggcgcca acagcgaatg gatgcattgc aacggcgcca tcggctacgg caatacgccc 1080
 tga 1083

<210> 218

<211> 360
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(29)

<400> 218
 Met Thr Phe Val Lys Thr Ile Thr Gly Arg Arg Ala Ile Ala Ala Phe
 1 5 10 15
 Leu Cys Leu Ala Gly Leu Tyr Met Ala Pro Ala Asn Ala Gln Thr Cys
 20 25 30
 Ile Thr Ser Ser Gln Thr Gly Thr Asn Asn Gly Asn Tyr Phe Ser Phe
 35 40 45
 Trp Lys Asp Ser Pro Gly Thr Val Asn Phe Cys Met Tyr Pro Asn Gly
 50 55 60
 Arg Tyr Thr Ser Asn Trp Ser Gly Ile Asn Asn Trp Val Gly Gly Lys
 65 70 75 80
 Gly Trp Ser Thr Gly Ser Ser Arg Thr Val Ser Tyr Ser Gly Ser Phe
 85 90 95
 Asn Ser Pro Gly Asn Gly Tyr Leu Thr Leu Tyr Gly Trp Thr Thr Asn
 100 105 110
 Pro Leu Ile Glu Tyr Tyr Ile Val Glu Asn Trp Gly Asn Tyr Arg Pro
 115 120 125
 Pro Gly Gly Gln Gly Tyr Met Gly Thr Val Asn Ser Asp Gly Ala Thr
 130 135 140
 Tyr Asp Ile Tyr Arg Thr Phe Arg Asp Asn Gln Pro Cys Ile Thr Gly
 145 150 155 160
 Asn Ser Cys Asp Phe Tyr Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg
 165 170 175
 Ser Ser Gly Thr Ile Thr Thr Ala Asn His Phe Ala Ala Trp Asn Ser
 180 185 190
 Leu Gly Met Asn Leu Gly Gln His Asn Tyr Gln Val Met Ala Thr Glu
 195 200 205
 Gly Tyr Gln Ser Ser Gly Ser Ser Asp Ile Thr Val Thr Glu Gly Gly
 210 215 220
 Gly Gly Ser Ser Asn Gly Gly Ser Ser Asn Gly Gly Ser Ser Asn Gly
 225 230 235 240
 Gly Ser Ser Asn Gly Gly Gly Gly Gly Thr Lys Ser Phe Thr Val Arg
 245 250 255
 Ala Arg Gly Thr Ala Gly Gly Glu Ser Ile Thr Leu Arg Val Asn Asn
 260 265 270
 Gln Asn Val Gln Thr Trp Thr Leu Gly Thr Gly Met Gln Asn Tyr Thr
 275 280 285
 Ala Ser Thr Ser Leu Ser Gly Ile Thr Val His Phe Thr Asn Asp
 290 295 300
 Gly Gly Ser Arg Asp Val Gln Val Asp Tyr Ile Gln Val Asn Gly Ser
 305 310 315 320
 Thr Arg Gln Ser Glu Ala Gln Ser Tyr Asn Thr Gly Ala Tyr Leu Asn
 325 330 335
 Gly Arg Cys Gly Gly Gly Asn Ser Glu Trp Met His Cys Asn Gly
 340 345 350
 Ala Ile Gly Tyr Gly Asn Thr Pro
 355 360

<210> 219
 <211> 1029
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 219
 atgacatcag gtctcaagaa agtgatggca ttctgtctgtc tcgccaccct tggcgtttcg
 gcgcgatgcc agacatgtat tcagtccagt cagaccggca ccaacaacgg attctatttc
 Page 161

60
 120

tccttctgga	aggacaaccc	gggcacggtg	cagttctgcc	tgcagagcgg	cggtcgttac	180
acctccaact	ggaacggcat	caacaactgg	gtgggcggca	aggggtggca	gaccggcgca	240
cggcgacgg	tgaactactc	gggctcgttc	aactcgccgg	gcaacggcta	tctggcgctg	300
tacggctgga	ccaccaatcc	gctggtcgag	tactacatcg	tcgacagctg	gggcagcttc	360
cgtccgccgg	gcaacactgc	aggcctgtgg	gtactgggtga	acagcgatgg	cggcacctac	420
gacatctatc	gcgcgcatcg	cagtaacgcg	ccctgcatca	ccggcagcag	ctgcgacttc	480
gaccagtact	ggagcgtgcg	acagtcgaag	cgcgtcggcg	gcaccatcac	caccggcaac	540
cacttcgatg	cctgggcgaa	ccaccagatg	aatctgggcc	agttcaacta	ccagatcatg	600
gtaccgagg	gtttccagag	caacggcagc	tccgacatca	ccgtcagtga	atgcaccagc	660
aattgcggcg	gtggcgggcg	cggcgggggg	ggcagcaaca	gcatcacggg	gcgcgcgcgc	720
ggcacgggcg	gcggcgagca	gatccggctg	cgggtgaaca	acaccacggg	gcaaacctgg	780
acgctgacca	ccagctacca	gaacttcacg	gcttcgacct	cgttgagcgg	cggcaccatc	840
gtcgagtact	tcaacgacag	ttccggccat	gacgtgcagg	tcgactacat	catcgtgaat	900
ggcgtgaccc	gccagtccga	atcgcagagc	tacaacaccg	ggctgtatgc	caacgggcgt	960
tgcggcgggcg	gctccaacag	cgaagtggatg	cattgcaacg	gtgccattgg	atacggaat	1020
accccgtaa						1029

<210> 220
 <211> 342
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(23)

<400> 220
 Met Thr Ser Gly Leu Lys Lys Val Met Ala Phe Val Cys Leu Ala Thr
 1 5 10 15
 Leu Gly Val Ser Ala His Ala Gln Thr Cys Ile Gln Ser Ser Gln Thr
 20 25 30
 Gly Thr Asn Gly Phe Tyr Phe Ser Phe Trp Lys Asp Asn Pro Gly
 35 40 45
 Thr Val Gln Phe Cys Leu Gln Ser Gly Gly Arg Tyr Thr Ser Asn Trp
 50 55 60
 Asn Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Gln Thr Gly Ala
 65 70 75 80
 Arg Arg Thr Val Asn Tyr Ser Gly Ser Phe Asn Ser Pro Gly Asn Gly
 85 90 95
 Tyr Leu Ala Leu Tyr Gly Trp Thr Thr Asn Pro Leu Val Glu Tyr Tyr
 100 105 110
 Ile Val Asp Ser Trp Gly Ser Phe Arg Pro Pro Gly Asn Thr Ala Gly
 115 120 125
 Leu Trp Val Leu Val Asn Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Arg
 130 135 140
 Ala His Arg Ser Asn Ala Pro Cys Ile Thr Gly Ser Ser Cys Asp Phe
 145 150 155 160
 Asp Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg Val Gly Gly Thr Ile
 165 170 175
 Thr Thr Gly Asn His Phe Asp Ala Trp Ala Asn His Gln Met Asn Leu
 180 185 190
 Gly Gln Phe Asn Tyr Gln Ile Met Ala Thr Glu Gly Phe Gln Ser Asn
 195 200 205
 Gly Ser Ser Asp Ile Thr Val Ser Glu Cys Thr Ser Asn Cys Gly Gly
 210 215 220
 Gly Gly Gly Gly Gly Gly Ser Asn Ser Ile Thr Val Arg Ala Arg
 225 230 235 240
 Gly Thr Gly Gly Gly Glu Gln Ile Arg Leu Arg Val Asn Asn Thr Thr
 245 250 255
 Val Gln Thr Trp Thr Leu Thr Thr Ser Tyr Gln Asn Phe Thr Ala Ser
 260 265 270
 Thr Ser Leu Ser Gly Gly Thr Ile Val Glu Tyr Phe Asn Asp Ser Ser
 275 280 285
 Gly His Asp Val Gln Val Asp Tyr Ile Ile Val Asn Gly Val Thr Arg
 290 295 300
 Gln Ser Glu Ser Gln Ser Tyr Asn Thr Gly Leu Tyr Ala Asn Gly Arg
 305 310 315 320

Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His Cys Asn Gly Ala Ile
 325 330 335
 Gly Tyr Gly Asn Thr Pro
 340

<210> 221
 <211> 1044
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 221
 atgattgtta gtttcaagag cgtgaaggca ctcgctgtcc tggccgtgct cggcgtgacc 60
 gccgcgcagg cgcaaacctg catcaattcc agccagaccg gcaccaacaa cggcaattat 120
 ttttcattct ggaaagacaa cccgggcacg gtgaccttct gcatgtatgc caacggccgc 180
 tacacctcca actggagcgg catcaacaac tgggtgggtg gcaagggctg gcagaccggc 240
 tcgaatcgca cggtgacctt ctccggttcg ttcaactcgc ccggcaacgg ctacctcacc 300
 ctgtacgggt ggaccacgaa tccgctgatc gagtactaca tcgtcgacag ttggggcagt 360
 tatcgaccgc ccggcggcca gggcttcatg ggacccgtga cgaccgacgg cggcacctac 420
 gacatctatc gcagcgacgg cgtgaaccag ccttccatca tcggcaccgc gacgttctac 480
 cagtactgga gcgtgcggca gtcgaagcgc gtggggggca ccatcaccac cgccaaccac 540
 ttcaatgcct gggcgacgct gggcatgaac ctggggccagc acaactacca ggtcatggcc 600
 accgaggggt accagagcag cggcagctcc gacatcaccg tgaccgaagg cggcggcagc 660
 tcgtcgtcgt cgagcggcgg cggcagcacc agcagcggcg gtggcggcag caagagcttc 720
 acggtgcgcg cccgcggcac ggtcggcggc gaaaacatcc agctgcaggc caacaaccag 780
 acggtggcga gctggaacct gaccaccagc atgcagaact acaacgcctc gaccagcctg 840
 agtggcggca tcaccgtggt ctacaccaac gacggcggta accgcgacgt ccaggtcgac 900
 tacatcaccg tgaacggcca gaccgcagc tccgaagcgc agagtttcaa caccgggctg 960
 tatgccaacg gacgttgtgg cggcggctcg aacagcgagt ggatgcattg caatggcgcg 1020
 atcggctacg gcaacacgcc gtaa 1044

<210> 222
 <211> 347
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(24)

<400> 222
 Met Ile Val Ser Phe Lys Ser Val Lys Ala Leu Ala Cys Leu Ala Val
 1 5 10 15
 Leu Gly Val Thr Ala Ala Gln Ala Gln Thr Cys Ile Asn Ser Ser Gln
 20 25 30
 Thr Gly Thr Asn Asn Gly Asn Tyr Phe Ser Phe Trp Lys Asp Asn Pro
 35 40 45
 Gly Thr Val Thr Phe Cys Met Tyr Ala Asn Gly Arg Tyr Thr Ser Asn
 50 55 60
 Trp Ser Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Gln Thr Gly
 65 70 75 80
 Ser Asn Arg Thr Val Thr Tyr Ser Gly Ser Phe Asn Ser Pro Gly Asn
 85 90 95
 Gly Tyr Leu Thr Leu Tyr Gly Trp Thr Thr Asn Pro Leu Ile Glu Tyr
 100 105 110
 Tyr Ile Val Asp Ser Trp Gly Ser Tyr Arg Pro Pro Gly Gly Gln Gly
 115 120 125
 Phe Met Gly Thr Val Thr Thr Asp Gly Gly Thr Tyr Asp Ile Tyr Arg
 130 135 140
 Thr Gln Arg Val Asn Gln Pro Ser Ile Ile Gly Thr Ala Thr Phe Tyr
 145 150 155 160
 Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg Val Gly Gly Thr Ile Thr
 165 170 175
 Thr Ala Asn His Phe Asn Ala Trp Ala Thr Leu Gly Met Asn Leu Gly
 180 185 190

Gln His Asn Tyr Gln Val Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly
 195 200 205
 Ser Ser Asp Ile Thr Val Thr Glu Gly Gly Gly Ser Ser Ser Ser
 210 215 220
 Ser Gly Gly Gly Ser Thr Ser Ser Gly Gly Gly Gly Ser Lys Ser Phe
 225 230 235 240
 Thr Val Arg Ala Arg Gly Thr Val Gly Gly Glu Asn Ile Gln Leu Gln
 245 250 255
 Val Asn Asn Gln Thr Val Ala Ser Trp Asn Leu Thr Thr Ser Met Gln
 260 265 270
 Asn Tyr Asn Ala Ser Thr Ser Leu Ser Gly Gly Ile Thr Val Val Tyr
 275 280 285
 Thr Asn Asp Gly Gly Asn Arg Asp Val Gln Val Asp Tyr Ile Thr Val
 290 295 300
 Asn Gly Gln Thr Arg Gln Ser Glu Ala Gln Ser Phe Asn Thr Gly Leu
 305 310 315 320
 Tyr Ala Asn Gly Arg Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His
 325 330 335
 Cys Asn Gly Ala Ile Gly Tyr Gly Asn Thr Pro
 340 345

<210> 223
 <211> 642
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 223
 atgtttaagt ttaaaaagaa tttcttagtt ggattatcgg cagctttaat gagtattagc 60
 ttgttttcgg caaccgcctc tgcagctagc acagactact ggcaaaattg gactgatggg 120
 ggcggtatag taaacgctgt caatgggtct ggcgggaatt acagtgttaa ttggtctaat 180
 accggaatt tcgttggttg taaagggttg actacaggtt cgccatttag gacgataaac 240
 tataatgccg gaggttgggc accgaatgga aatggatatt taactttata tggttggacg 300
 agatcacctc tcatagaata ttatgtagtg gattcatggg gtacttatag acctactgga 360
 acgtataaag gtactgtaaa aagtgatggg ggtacatatg acatatatac aactacacgt 420
 tataacgcac cticcattga tggcgatcgc actactttta cgcagtactg gagtgttcgc 480
 caaacgaaga gaccaaccgg aagcaacgct acaatcactt tcagcaatca tgttaacgca 540
 tggaagagcc atggaatgaa tctgggcagt aattgggcctt accaagtcac ggcgacagaa 600
 ggatatcaaa gtagtggaag ttctaacgta acagtgtggt aa 642

<210> 224
 <211> 213
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 224
 Met Phe Lys Phe Lys Lys Asn Phe Leu Val Gly Leu Ser Ala Ala Leu
 1 5 10 15
 Met Ser Ile Ser Leu Phe Ser Ala Thr Ala Ser Ala Ala Ser Thr Asp
 20 25 30
 Tyr Trp Gln Asn Trp Thr Asp Gly Gly Gly Ile Val Asn Ala Val Asn
 35 40 45
 Gly Ser Gly Gly Asn Tyr Ser Val Asn Trp Ser Asn Thr Gly Asn Phe
 50 55 60
 Val Val Gly Lys Gly Trp Thr Thr Gly Ser Pro Phe Arg Thr Ile Asn
 65 70 75 80
 Tyr Asn Ala Gly Val Trp Ala Pro Asn Gly Asn Gly Tyr Leu Thr Leu
 85 90 95
 Tyr Gly Trp Thr Arg Ser Pro Leu Ile Glu Tyr Tyr Val Val Asp Ser
 100 105 110
 Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys Gly Thr Val Lys Ser

115 120 125
 Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr Thr Arg Tyr Asn Ala Pro
 130 135 140
 Ser Ile Asp Gly Asp Arg Thr Thr Phe Thr Gln Tyr Trp Ser Val Arg
 145 150 155 160
 Gln Thr Lys Arg Pro Thr Gly Ser Asn Ala Thr Ile Thr Phe Ser Asn
 165 170 175
 His Val Asn Ala Trp Lys Ser His Gly Met Asn Leu Gly Ser Asn Trp
 180 185 190
 Ala Tyr Gln Val Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly Ser Ser
 195 200 205
 Asn Val Thr Val Trp
 210

<210> 225
 <211> 1059
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 225
 atgtttgtta gtctcaggaa gacggccttg gcgtgcctgt tgctcgccgg cctcggaatc 60
 tcgattcac aagcccagac ctgcatcac tccagcggga cgggcaccaa caacggccac 120
 tactattcct tctggaagga cagtggcggc accgtcaact tctgcatgta cgcgaaacggc 180
 cgctacacct ccaactggag cggcatcaac aactgggtgg gcggcaaggg ctggcagacc 240
 ggctcacgcc ggacgatcag ctactcgggc tcgttcaact cacccggaac tggttatctc 300
 accctgtacg gttggaccac caatccattg atcgagtact acatcgctga caactggggc 360
 acgtaccggc cgccgggagg ctccgggctac atgggcacag tgacgagcga cggcggcacc 420
 tacgacgtct atcgaccca gcgcgtaaac cagccttcca tcatcggcac cgcgacgttc 480
 tatcaatact ggagcgtgcg ccagcagaag cggaccggcg ggaccatcac caccggcaat 540
 cacttcgacg cctggggccgc atacggaatg aacctcggca cccacaacta ccagatcatg 600
 gcgaccgagg gttaccagag cagcggcagt tcggacatca cggtgagcga gggcgggtggc 660
 agttcacaga gcagcagtc gtcgagcagc agcagttcgt cctcttcgag cggcggcggc 720
 ggcacgaaga gcttcacggt ccgcgcgcgc ggcacggcgg gcggtgaatc catcacgctg 780
 cgctgaaca accagaacgt gcagacctgg acgctgggca cgtcgatgca gaactacacc 840
 gcatcgacca cgctctccgg tggcatcacc gtgcgtgata ccaacgacag cggcaatcgc 900
 gacgtgcagg tggactacat cgtcgtgaac ggcgccacc gccagtccga ggcgcagagc 960
 tacaacaccg gtctctatgc caacgggtcgt tgcggcggcg gctccaacag cgagtggatg 1020
 cactgcaacg ggcagatcgg ctacgggaat actccctag 1059

<210> 226
 <211> 352
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(25)

<400> 226
 Met Phe Val Ser Leu Arg Lys Thr Ala Trp Ala Cys Leu Leu Leu Ala
 1 5 10 15
 Gly Leu Gly Ile Ser Thr Ser Gln Ala Gln Thr Cys Ile Thr Ser Ser
 20 25 30
 Gly Thr Gly Thr Asn Asn Gly His Tyr Tyr Ser Phe Trp Lys Asp Ser
 35 40 45
 Gly Gly Thr Val Asn Phe Cys Met Tyr Ala Asn Gly Arg Tyr Thr Ser
 50 55 60
 Asn Trp Ser Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Gln Thr
 65 70 75 80
 Gly Ser Arg Arg Thr Ile Ser Tyr Ser Gly Ser Phe Asn Ser Pro Gly
 85 90 95
 Asn Gly Tyr Leu Thr Leu Tyr Gly Trp Thr Thr Asn Pro Leu Ile Glu
 100 105 110
 Tyr Tyr Ile Val Asp Asn Trp Gly Thr Tyr Arg Pro Pro Gly Gly Ser

115 120 125
 Gly Tyr Met Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Val Tyr
 130 135 140
 Arg Thr Gln Arg Val Asn Gln Pro Ser Ile Ile Gly Thr Ala Thr Phe
 145 150 155 160
 Tyr Gln Tyr Trp Ser Val Arg Gln Gln Lys Arg Thr Gly Gly Thr Ile
 165 170 175
 Thr Thr Gly Asn His Phe Asp Ala Trp Ala Ala Tyr Gly Met Asn Leu
 180 185 190
 Gly Thr His Asn Tyr Gln Ile Met Ala Thr Glu Gly Tyr Gln Ser Ser
 195 200 205
 Gly Ser Ser Asp Ile Thr Val Ser Glu Gly Gly Gly Ser Ser Ser Ser
 210 215 220
 Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Gly Gly Gly
 225 230 235 240
 Gly Thr Lys Ser Phe Thr Val Arg Ala Arg Gly Thr Ala Gly Gly Glu
 245 250 255
 Ser Ile Thr Leu Arg Val Asn Asn Gln Asn Val Gln Thr Trp Thr Leu
 260 265 270
 Gly Thr Ser Met Gln Asn Tyr Thr Ala Ser Thr Thr Leu Ser Gly Gly
 275 280 285
 Ile Thr Val Ala Tyr Thr Asn Asp Ser Gly Asn Arg Asp Val Gln Val
 290 295 300
 Asp Tyr Ile Val Val Asn Gly Ala Thr Arg Gln Ser Glu Ala Gln Ser
 305 310 315 320
 Tyr Asn Thr Gly Leu Tyr Ala Asn Gly Arg Cys Gly Gly Gly Ser Asn
 325 330 335
 Ser Glu Trp Met His Cys Asn Gly Gln Ile Gly Tyr Gly Asn Thr Pro
 340 345 350

<210> 227
 <211> 747
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 227
 atgggcggca cgactggtag tggcggctca gccgccgccg gcgcaggcac gagtggaagc 60
 gcgggcggta ccgccggagc gctcggcccc ggccgtaccc agggcagcgg tggcgcagcc 120
 ggtggtacga gcggaacggg cggggccatc agcagcagct gcacggaagc tgacaagacg 180
 gtctgcaaca acgaaaccgg tcgccactgc aattacacgt acgagtattg gaaggaccag 240
 ggaagcggtg gcctcgtgaa caaagccgac ggcttcagcg tcaactggaa caacatcaac 300
 aatctgctgg gtgcgaagg tctgaggccc ggatcgtcga atcagacggt gacctaccag 360
 gcaactacc agccgaacgg caattcatal ctgtgcgtat atggatggac gcaaaacccc 420
 ctctgcgaat actacatcgt cgatagctgg ggcagctggc gcccgcgggg gggaacgtcc 480
 atgggcaccg tcaacgcgga cggcggcacc tacgacatct accgcaccca gcgcgtcaac 540
 cagccttcca tcgaaggcac caagaccttc tatcaatact ggagcggttc cactcagaag 600
 cgacagagcg gaacgatcac ggttgccgct cacttcgacg cctgggcgac gaaggggatg 660
 aacatgggga gtctgtacga ggtgtcgatg accgtcgagg gctatcaaag cagcgggacc 720
 gccgacgtga gcttctcgat gaagtga 747

<210> 228
 <211> 248
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(39)

<400> 228
 Met Gly Gly Thr Thr Gly Ser Gly Gly Ser Ala Ala Ala Gly Ala Gly
 1 5 10 15
 Thr Ser Gly Ser Ala Gly Gly Thr Ala Gly Ala Leu Gly Pro Gly Gly
 20 25 30

Thr Gln Gly Ser Gly Gly Ala Ala Gly Gly Thr Ser Gly Thr Gly Gly
 35 40 45
 Ala Ile Ser Ser Ser Cys Thr Glu Ala Asp Lys Thr Val Cys Asn Asn
 50 55 60
 Glu Thr Gly Arg His Cys Asn Tyr Thr Tyr Glu Tyr Trp Lys Asp Gln
 65 70 75 80
 Gly Ser Gly Cys Leu Val Asn Lys Ala Asp Gly Phe Ser Val Asn Trp
 85 90 95
 Asn Asn Ile Asn Asn Leu Leu Gly Arg Lys Gly Leu Arg Pro Gly Ser
 100 105 110
 Ser Asn Gln Thr Val Thr Tyr Gln Ala Asn Tyr Gln Pro Asn Gly Asn
 115 120 125
 Ser Tyr Leu Cys Val Tyr Gly Trp Thr Gln Asn Pro Leu Val Glu Tyr
 130 135 140
 Tyr Ile Val Asp Ser Trp Gly Ser Trp Arg Pro Pro Gly Gly Thr Ser
 145 150 155 160
 Met Gly Thr Val Asn Ala Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Thr
 165 170 175
 Gln Arg Val Asn Gln Pro Ser Ile Glu Gly Thr Lys Thr Phe Tyr Gln
 180 185 190
 Tyr Trp Ser Val Arg Thr Gln Lys Arg Thr Ser Gly Thr Ile Thr Val
 195 200 205
 Ala Ala His Phe Asp Ala Trp Ala Thr Lys Gly Met Asn Met Gly Ser
 210 215 220
 Leu Tyr Glu Val Ser Met Thr Val Glu Gly Tyr Gln Ser Ser Gly Thr
 225 230 235 240
 Ala Asp Val Ser Phe Ser Met Lys
 245

<210> 229
 <211> 642
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 229
 atgtttaagt ttacaaagaa attccttagtt ggggttaacgg cagctttgat gagtattagc 60
 ttgttttcgg caaacgcctc tgcagctaac acagactact ggcaaaaattg gactgatggg 120
 ggcggaacag taaacgctgt caatgggtct ggcgggaatt acagtgtgaa ttggtctaatt 180
 accgggaatt tcgttggttg taaaggttgg actacaggtt cgccatttag gacgataaac 240
 tataatgccg gagtttgggc gccgaatggc aatgcatatt tgactttata tggttggacg 300
 cgatcacccc tcatagaata ttatgtagtg gattcatggg gtacttatag acctactgga 360
 acgtataaag gtacggttta cagtgtatggg ggtacatatg acgtgtacac aactacacgt 420
 tatgatgcac cticcattga tggcgataaa actactttta cgcagtactg gagtgttcgc 480
 cagtcgaaga gaccaactgg aagcaacgct acaatcactt tcagcaatca cgttaacgca 540
 tgggaagat atgggatgaa tctgggtagt aattggtcct accaagtctt agcgacagag 600
 ggatatcaaa gtagtggaag ttctaacgta acagtgtggt aa 642

<210> 230
 <211> 213
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 230
 Met Phe Lys Phe Thr Lys Lys Phe Leu Val Gly Leu Thr Ala Ala Leu
 1 5 10 15
 Met Ser Ile Ser Leu Phe Ser Ala Asn Ala Ser Ala Ala Asn Thr Asp
 20 25 30
 Tyr Trp Gln Asn Trp Thr Asp Gly Gly Thr Val Asn Ala Val Asn
 35 40 45
 Gly Ser Gly Gly Asn Tyr Ser Val Asn Trp Ser Asn Thr Gly Asn Phe

50 55 60
 Val Val Gly Lys Gly Trp Thr Thr Gly Ser Pro Phe Arg Thr Ile Asn
 65 70 75 80
 Tyr Asn Ala Gly Val Trp Ala Pro Asn Gly Asn Ala Tyr Leu Thr Leu
 85 90 95
 Tyr Gly Trp Thr Arg Ser Pro Leu Ile Glu Tyr Tyr Val Val Asp Ser
 100 105 110
 Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys Gly Thr Val Tyr Ser
 115 120 125
 Asp Gly Gly Thr Tyr Asp Val Tyr Thr Thr Thr Arg Tyr Asp Ala Pro
 130 135 140
 Ser Ile Asp Gly Asp Lys Thr Thr Phe Thr Gln Tyr Trp Ser Val Arg
 145 150 155 160
 Gln Ser Lys Arg Pro Thr Gly Ser Asn Ala Thr Ile Thr Phe Ser Asn
 165 170 175
 His Val Asn Ala Trp Lys Arg Tyr Gly Met Asn Leu Gly Ser Asn Trp
 180 185 190
 Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln Ser Ser Gly Ser Ser
 195 200 205
 Asn Val Thr Val Trp
 210

<210> 231
 <211> 1008
 <212> DNA
 <213> Bacteria

<400> 231
 atgaacctgc tcgtccagcc gaggcgtcgc agacgcggtc cggtcacctt gctcgtcagg 60
 agcgcgtggg ccgtcgcgct ggccgcgctc gccgcgtga tgctgccggg caccgcccag 120
 gccgacacgg tcgtcacgac caaccaggag ggccaacaac acggctacta ctactcgttc 180
 tggaccgaca gccagggcac cgtctccatg aacatgggct ccggcgggtca gtacagcacc 240
 tcgtggcgca acaccggcaa cttcgtcgcg ggcaagggct gggccaacgg cgcccgccgg 300
 accgtgcagt actcgggcag cttcaacccc tccggcaacg cgtacctggc gctctacgga 360
 tggacgtcga acccgctcgt cgagtactac atcgtcgaca actggggcac ctaccggccc 420
 acgggcgagt acaagggcac cgtcaccagc gacggcgcca cctacgacat ctacaagacg 480
 acccgcgctca acaagccctc cgtcaggggc acccgcacct tcgaccagta ctggagcgctc 540
 cggcaggcga agcggaccgg cggcaccatc acgaccggca accacttcga cgcgtggggc 600
 cgggcccggga tggccgctcg caacttcagc tactacatga tcatggccac cgagggctac 660
 cagagcagcg gcagctccag catcaacgtc ggccgggacc gccgcggcga caacggcggc 720
 ggcgacaacg ggggcggtgg cggcggtgac accgccacgg tgtccgcccg gcagaagtgg 780
 ggcgaccggt acaacctcga cgtctccgtc agcggcgcca gcgactggac ggtgacgatg 840
 aacgtgcccgt ccccgccgaa ggtcctgtcg acctggaacg tcaacgccag ctatcccagt 900
 gcgcagacgc tggaccgccag gtcgaacggc agcggcaaca actggggcgc caccatccag 960
 gccaacggca actggacctg gccagcgctg tcctgcagcg cgggctga 1008

<210> 232
 <211> 335
 <212> PRT
 <213> Bacteria

<220>
 <221> SIGNAL
 <222> (1)...(41)

<400> 232
 Met Asn Leu Leu Val Gln Pro Arg Arg Arg Arg Gly Pro Val Thr
 1 5 10 15
 Leu Leu Val Arg Ser Ala Trp Ala Val Ala Leu Ala Ala Leu Ala Ala
 20 25 30
 Leu Met Leu Pro Gly Thr Ala Gln Ala Asp Thr Val Val Thr Thr Asn
 35 40 45
 Gln Glu Gly Thr Asn Asn Gly Tyr Tyr Tyr Ser Phe Trp Thr Asp Ser
 50 55 60
 Gln Gly Thr Val Ser Met Asn Met Gly Ser Gly Gln Tyr Ser Thr
 65 70 75 80
 Ser Trp Arg Asn Thr Gly Asn Phe Val Ala Gly Lys Gly Trp Ala Asn
 85 90 95
 Gly Gly Arg Arg Thr Val Gln Tyr Ser Gly Ser Phe Asn Pro Ser Gly

100 105 110
 Asn Ala Tyr Leu Ala Leu Tyr Gly Trp Thr Ser Asn Pro Leu Val Glu
 115 120 125
 Tyr Tyr Ile Val Asp Asn Trp Gly Thr Tyr Arg Pro Thr Gly Glu Tyr
 130 135 140
 Lys Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Lys Thr
 145 150 155 160
 Thr Arg Val Asn Lys Pro Ser Val Glu Gly Thr Arg Thr Phe Asp Gln
 165 170 175
 Tyr Trp Ser Val Arg Gln Ala Lys Arg Thr Gly Gly Thr Ile Thr Thr
 180 185 190
 Gly Asn His Phe Asp Ala Trp Ala Arg Ala Gly Met Pro Leu Gly Asn
 195 200 205
 Phe Ser Tyr Tyr Met Ile Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly
 210 215 220
 Ser Ser Ser Ile Asn Val Gly Gly Thr Gly Arg Gly Asp Asn Gly Gly
 225 230 235 240
 Gly Asp Asn Gly Gly Gly Gly Gly Cys Thr Ala Thr Val Ser Ala
 245 250 255
 Gly Gln Lys Trp Gly Asp Arg Tyr Asn Leu Asp Val Ser Val Ser Gly
 260 265 270
 Ala Ser Asp Trp Thr Val Thr Met Asn Val Pro Ser Pro Ala Lys Val
 275 280 285
 Leu Ser Thr Trp Asn Val Asn Ala Ser Tyr Pro Ser Ala Gln Thr Leu
 290 295 300
 Thr Ala Arg Ser Asn Gly Ser Gly Asn Asn Trp Gly Ala Thr Ile Gln
 305 310 315 320
 Ala Asn Gly Asn Trp Thr Trp Pro Ser Val Ser Cys Ser Ala Gly
 325 330 335

<210> 233
 <211> 1071
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 233
 atgtctatgt ttttgagtct caaaagagtg gcggcgctcg tctgcgtcgc aggggtttggc 60
 atttcggcgg cgaacgctca gtgcgtcact tcgagccaga caggaaccaa caacgggttc 120
 tatttttcgt tctggaaaga tagtccggga accgtgaatt tctgcaacca gagcgggtggc 180
 cgctacacat ccaattggag cggatatcaac aactgggtcg gtggcaagggt ttggcagacc 240
 ggctcgcgaa ggtcgtgag ctactccggt tcgttcaatt cgccgggcaa cgggtatctg 300
 accctctatg ggtggaccac caatccgctc atcgagtact acatcgtcga caactggggc 360
 tcgtatcgcc cgcggggcgg acaggggttc atgggcacgg tgaccagcga cggcggcacg 420
 tacgatgtct accgcacaca gcgcgtcaat caaccctgca tcaccggcag cagttgcacc 480
 ttctatcaat actggagcgt gcggcagtcg aagagaacgg gcggcacgat caccgacggc 540
 aatcacattg acgcgtgggc gagttacggc atgaacctgg gcgctcacia ctaccagatc 600
 atggcgaccg agggttatca aagcagcggg agctctgaca tcacggtcag tgaaggcagc 660
 agcagtagca gcagtagcag cagttcgagc agtagctcga gcagcagctc cagcagcagc 720
 agcggcggcg gtggcaccaa gagcttcacg gtccgcgcgc gcggcggtggc cggcggggaa 780
 tccatcacgt tgcgcgtgaa caatcagaac gtgcagacct ggactctcgg caccggcatg 840
 cagaactaca cggcgtcgac gtctttgagt ggcgccatca cgggtgcgta taccaacgat 900
 ggcggcagtc gcgacgtgca ggttgactac atcatcgtga acggccagac gcgtcagtcg 960
 gaagcgcaga gctacaacac cgggctttat gccaacggcc gttgcggtgg cggcggcaac 1020
 agcgaatgga tgcattgcaa tggcgccatt ggctacggga acacgccgta g 1071

<210> 234
 <211> 356
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(26)

<400> 234
 Met Ser Met Phe Leu Ser Leu Lys Arg Val Ala Ala Leu Val Cys Val
 1 5 10 15
 Ala Gly Phe Gly Ile Ser Ala Ala Asn Ala Gln Cys Val Thr Ser Ser
 20 25 30
 Gln Thr Gly Thr Asn Asn Gly Phe Tyr Phe Ser Phe Trp Lys Asp Ser
 35 40 45
 Pro Gly Thr Val Asn Phe Cys Asn Gln Ser Gly Gly Arg Tyr Thr Ser
 50 55 60
 Asn Trp Ser Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Gln Thr
 65 70 75 80
 Gly Ser Arg Arg Val Ser Tyr Ser Gly Ser Phe Asn Ser Pro Gly
 85 90 95
 Asn Gly Tyr Leu Thr Leu Tyr Gly Trp Thr Thr Asn Pro Leu Ile Glu
 100 105 110
 Tyr Tyr Ile Val Asp Asn Trp Gly Ser Tyr Arg Pro Pro Gly Gly Gln
 115 120 125
 Gly Phe Met Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Val Tyr
 130 135 140
 Arg Thr Gln Arg Val Asn Gln Pro Cys Ile Thr Gly Ser Ser Cys Thr
 145 150 155 160
 Phe Tyr Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg Thr Gly Gly Thr
 165 170 175
 Ile Thr Thr Gly Asn His Phe Asp Ala Trp Ala Ser Tyr Gly Met Asn
 180 185 190
 Leu Gly Ala His Asn Tyr Gln Ile Met Ala Thr Glu Gly Tyr Gln Ser
 195 200 205
 Ser Gly Ser Ser Asp Ile Thr Val Ser Glu Gly Ser Ser Ser Ser Ser
 210 215 220
 Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
 225 230 235 240
 Ser Gly Gly Gly Gly Thr Lys Ser Phe Thr Val Arg Ala Arg Gly Val
 245 250 255
 Ala Gly Gly Glu Ser Ile Thr Leu Arg Val Asn Asn Gln Asn Val Gln
 260 265 270
 Thr Trp Thr Leu Gly Thr Gly Met Gln Asn Tyr Thr Ala Ser Thr Ser
 275 280 285
 Leu Ser Gly Gly Ile Thr Val Ala Tyr Thr Asn Asp Gly Gly Ser Arg
 290 295 300
 Asp Val Gln Val Asp Tyr Ile Ile Val Asn Gly Gln Thr Arg Gln Ser
 305 310 315 320
 Glu Ala Gln Ser Tyr Asn Thr Gly Leu Tyr Ala Asn Gly Arg Cys Gly
 325 330 335
 Gly Gly Gly Asn Ser Glu Trp Met His Cys Asn Gly Ala Ile Gly Tyr
 340 345 350
 Gly Asn Thr Pro
 355

<210> 235
 <211> 1539
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 235
 atgtcgaata acagatttgt gctgaatcgt gttgctgcag gtttgctgct gggtttctcg 60
 ctgctgtcat cagcagccat cgcccagaat gtggtggtaa atccttctac ggtccatcag 120
 accgtgcgcg gctttggcgg catgaacgcg ccgggctgga ttgatgacct taccaccgcc 180
 cagggtcaata aggcctatgg cagtggcgat ggccaggctg ggctctccat catgcgcatg 240
 cgcattgatc cgaactcggc agcctggaat atccaggctg cggctgccaag gcgggccaag 300
 gagctgggtg cgatcctgtt tgccacgccc tggctgcccgc ccgctacat gaaatccaac 360
 aaaagcctga ataacggcgg caagctgctg cccgagtatt acagcgccta caccaccac 420
 ctgctggatt ttgcgagttt catgtcgcgc aacggcgcac cgctgtatgc gatttcaatc 480
 cagaacgaac cggactggct gccggattat gagtcgtgtg cctggactgg tactgatttc 540
 gtcaattatc tgaataccca gggctcgcgt tttgggtgac tgaaagtgat tgcgccggaa 600
 tccttgggtt tcacgacctc gtattccgac cccatcctca acagcgccac ggcagcgccg 660
 catgtcgaca tcatcggcgg ccacctctac ggcgtgctgc ccaaggacta cccgctggcg 720

cgccagaagg	gcaaggaaat	ctggatgacc	gagcattaca	ccgagagcaa	gaactcgggt	780
gatgcctggc	cgctggcgct	ggacgtaggc	accgagctgc	accagagcat	ggtggccaac	840
tacaacgcct	acgtgtggtg	gtatgtgctg	cgagctacg	gcctgctgct	ggagaacggc	900
aatgtgagca	agcgcggtta	catcatgtcg	cagtacgcac	gcttcgtccg	ccccggctcc	960
aagcgcatcg	gcgcgacgga	aaagccgcac	gccgacgtgg	cggtgacggc	ctacaagacg	1020
ccggataacc	gcattgtgct	ggtggcggtg	aataccggtg	cgggcgaccg	tcagctgaac	1080
atcacggtgc	cgagcggcag	cgtgggttct	ttcagcaagt	tctccacttc	cggcacgctg	1140
aatgtgggca	gtggtggcag	ctacaaggct	aacaacggcg	cggtgagcct	gtacatcgat	1200
ccgcagagcg	tggccacgct	ggtgggtgat	ctgccgggca	cggcctccag	ctcttcggcg	1260
gcgtcctcgt	cctcttccag	tgacgccagc	tctgcttcga	gcagtgctag	cggcgcaccg	1320
gccctgtctg	gcagcagcga	ttaccccacg	ggcttcagca	agtgcgctga	tctgggtggt	1380
acctgtgccg	tgcttccggg	ctcgggctgg	acggcttctg	ggcgcaaggg	caagtgggtt	1440
gccaaagtacg	tcggtgtggg	caagagcatt	gcctgcacgg	tgacggcttt	cggcagcgat	1500
cccggtgggtg	cacccaacaa	gtgttcttac	cagaagtaa			1539

<210> 236

<211> 512

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample

<221> SIGNAL

<222> (1)...(28)

<400> 236

Met	Ser	Asn	Asn	Arg	Phe	Val	Leu	Asn	Arg	Val	Ala	Ala	Gly	Leu	Leu
1				5					10					15	
Leu	Gly	Phe	Ser	Leu	Leu	Ser	Ser	Ala	Ile	Ala	Gln	Asn	Val	Val	
			20					25				30			
Val	Asn	Pro	Ser	Thr	Val	His	Gln	Thr	Val	Arg	Gly	Phe	Gly	Gly	Met
		35				40					45				
Asn	Ala	Pro	Gly	Trp	Ile	Asp	Asp	Leu	Thr	Thr	Ala	Gln	Val	Asn	Lys
		50				55					60				
Ala	Tyr	Gly	Ser	Gly	Asp	Gly	Gln	Val	Gly	Leu	Ser	Ile	Met	Arg	Met
65					70				75					80	
Arg	Ile	Asp	Pro	Asn	Ser	Ala	Ala	Trp	Asn	Ile	Gln	Val	Pro	Ala	Ala
			85						90					95	
Lys	Arg	Ala	Lys	Glu	Leu	Gly	Ala	Ile	Leu	Phe	Ala	Thr	Pro	Trp	Ser
			100					105					110		
Pro	Pro	Ala	Tyr	Met	Lys	Ser	Asn	Lys	Ser	Leu	Asn	Asn	Gly	Gly	Lys
		115					120					125			
Leu	Leu	Pro	Glu	Tyr	Tyr	Ser	Ala	Tyr	Thr	Thr	His	Leu	Leu	Asp	Phe
		130				135					140				
Ala	Ser	Phe	Met	Ser	Arg	Asn	Gly	Ala	Pro	Leu	Tyr	Ala	Ile	Ser	Ile
145					150					155				160	
Gln	Asn	Glu	Pro	Asp	Trp	Leu	Pro	Asp	Tyr	Glu	Ser	Cys	Ala	Trp	Thr
			165						170					175	
Gly	Thr	Asp	Phe	Val	Asn	Tyr	Leu	Asn	Thr	Gln	Gly	Ser	Arg	Phe	Gly
		180						185					190		
Asp	Leu	Lys	Val	Ile	Ala	Pro	Glu	Ser	Leu	Gly	Phe	Thr	Thr	Ser	Tyr
		195					200					205			
Ser	Asp	Pro	Ile	Leu	Asn	Ser	Ala	Thr	Ala	Ala	Pro	His	Val	Asp	Ile
		210				215					220				
Ile	Gly	Gly	His	Leu	Tyr	Gly	Val	Leu	Pro	Lys	Asp	Tyr	Pro	Leu	Ala
225					230					235				240	
Arg	Gln	Lys	Gly	Lys	Glu	Ile	Trp	Met	Thr	Glu	His	Tyr	Thr	Glu	Ser
			245						250					255	
Lys	Asn	Ser	Gly	Asp	Ala	Trp	Pro	Leu	Ala	Leu	Asp	Val	Gly	Thr	Glu
			260					265					270		
Leu	His	Gln	Ser	Met	Val	Ala	Asn	Tyr	Asn	Ala	Tyr	Val	Trp	Trp	Tyr
		275					280					285			
Val	Arg	Arg	Ser	Tyr	Gly	Leu	Leu	Leu	Glu	Asn	Gly	Asn	Val	Ser	Lys
		290				295					300				
Arg	Gly	Tyr	Ile	Met	Ser	Gln	Tyr	Ala	Arg	Phe	Val	Arg	Pro	Gly	Ser
					310					315				320	
Lys	Arg	Ile	Gly	Ala	Thr	Glu	Lys	Pro	His	Ala	Asp	Val	Ala	Val	Thr
				325					330					335	

Ala Tyr Lys Thr Pro Asp Asn Arg Ile Val Leu Val Ala Val Asn Thr
 340 345 350
 Gly Ala Ala His Arg Gln Leu Asn Ile Thr Val Pro Ser Gly Ser Val
 355 360 365
 Gly Ser Phe Ser Lys Phe Ser Thr Ser Gly Thr Leu Asn Val Gly Ser
 370 375 380
 Gly Gly Ser Tyr Lys Val Asn Asn Gly Ala Val Ser Leu Tyr Ile Asp
 385 390 395 400
 Pro Gln Ser Val Ala Thr Leu Val Gly Asp Leu Pro Gly Thr Ala Ser
 405 410 415
 Ser Ser Ser Ala Ala Ser Ser Ser Ser Ser Ala Ala Ser Ser Ala
 420 425 430
 Ser Ser Ser Ala Ser Gly Ala Pro Ala Leu Ser Gly Ser Ser Asp Tyr
 435 440 445
 Pro Thr Gly Phe Ser Lys Cys Ala Asp Leu Gly Gly Thr Cys Ala Val
 450 455 460
 Pro Ser Gly Ser Gly Trp Thr Ala Phe Gly Arg Lys Gly Lys Trp Val
 465 470 475 480
 Ala Lys Tyr Val Gly Val Gly Lys Ser Ile Ala Cys Thr Val Thr Ala
 485 490 495
 Phe Gly Ser Asp Pro Gly Gly Ala Pro Asn Lys Cys Ser Tyr Gln Lys
 500 505 510

<210> 237
 <211> 1269
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 237
 atgattccac gcataaaaaa aacaatttgt gtactatttag tatgtttcac tatgctgtca 60
 gtcattgttag ggccaggcgc tactgaagtt ttggcagcaa gtgatgtaac agttaatgta 120
 tctgcagaga aacaagtgat tcgcggtttt ggagggatga atcatccggc ttgggctggg 180
 gatcttacag cagctcaaag agaaactgct tttggcaatg gacagaacca gttaggattt 240
 tcaatcttaa gaattcatgt agatgaaaat cgaaataatt ggtataaaga ggtggagact 300
 gcaaagagtg cggatcaaaca cggagcaatc gtttttgctt ctccttggaa tcctccaagt 360
 gatattggtg agacctttta tcggaatggt gacacatcgg ctaaacggct gaaatacaac 420
 aagtacgcag catacgcgca gcatcttaac gattttgtta ccttcatgaa gaataatggt 480
 gtgaatcttt acgcgatttc ggtccaaaac gagcctgatt acgctcacga gtggacgtgg 540
 tggacgccgc aagaaatact tcgctttatg agagaaaacg ccggctcgat caatgccgc 600
 gtcattgcgc ctgagtcatt tcaatacttg aagaatttgt cggacccgat cttgaacgat 660
 ccgaggctc ttgccaatat ggatattctc ggaactcacc tgtacggcac ccaggtcagc 720
 caattccctt atcctctttt caaacaataa ggagcgggga aggaaccttg gatgacggaa 780
 gtatactatc caaacagtga taccaactcg gcggatcgat ggcctgaggc attggatgtt 840
 tcacagcata ttcacaatgc gatggttagg ggggactttc aagcttatgt atggtgttac 900
 atccgaagat catatggagc tatgaaagaa gatggtacga tcagcaaacg cggctacaat 960
 atgggtcatt tctcaaagtt tgtgcgtccc ggctattgtaa ggattgatgc aacgaaaaac 1020
 cctaatacga acgtttacgt gtcagcctat aaaggtgaca acaaggctcg tattgttgcc 1080
 atcaataaaa gcaacacagg agtcaaccaa aactttgttt tgcagaatgg atctgcttca 1140
 aacgtatcta gatggatcac gagcagcagc agcaatctac aacctggaac gaatctcact 1200
 gtatcaggca atcatttttg ggctcatctt ccagctcaaa gcgtgacaac atttgttgta 1260
 aatcgtaa 1269

<210> 238
 <211> 422
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(32)

<400> 238
 Met Ile Pro Arg Ile Lys Lys Thr Ile Cys Val Leu Leu Val Cys Phe
 1 5 10 15

Thr Met Leu Ser Val Met Leu Gly Pro Gly Ala Thr Glu Val Leu Ala
 Ala Ser Asp Val Thr Val Asn Val Ser Ala Glu Lys Gln Val Ile Arg
 Gly Phe Gly Gly Met Asn His Pro Ala Trp Ala Gly Asp Leu Thr Ala
 Ala Gln Arg Glu Thr Ala Phe Gly Asn Gly Gln Asn Gln Leu Gly Phe
 Ser Ile Leu Arg Ile His Val Asp Glu Asn Arg Asn Asn Trp Tyr Lys
 Glu Val Glu Thr Ala Lys Ser Ala Val Lys His Gly Ala Ile Val Phe
 Ala Ser Pro Trp Asn Pro Pro Ser Asp Met Val Glu Thr Phe Asn Arg
 Asn Gly Asp Thr Ser Ala Lys Arg Leu Lys Tyr Asn Lys Tyr Ala Ala
 Tyr Ala Gln His Leu Asn Asp Phe Val Thr Phe Met Lys Asn Asn Gly
 Val Asn Leu Tyr Ala Ile Ser Val Gln Asn Glu Pro Asp Tyr Ala His
 Glu Trp Thr Trp Thr Pro Gln Glu Ile Leu Arg Phe Met Arg Glu
 Asn Ala Gly Ser Ile Asn Ala Arg Val Ile Ala Pro Glu Ser Phe Gln
 Tyr Leu Lys Asn Leu Ser Asp Pro Ile Leu Asn Asp Pro Gln Ala Leu
 Ala Asn Met Asp Ile Leu Gly Thr His Leu Tyr Gly Thr Gln Val Ser
 Gln Phe Pro Tyr Pro Leu Phe Lys Gln Lys Gly Ala Gly Lys Asp Leu
 Trp Met Thr Glu Val Tyr Tyr Pro Asn Ser Asp Thr Asn Ser Ala Asp
 Arg Trp Pro Glu Ala Leu Asp Val Ser Gln His Ile His Asn Ala Met
 Val Glu Gly Asp Phe Gln Ala Tyr Val Trp Trp Tyr Ile Arg Arg Ser
 Tyr Gly Pro Met Lys Glu Asp Gly Thr Ile Ser Lys Arg Gly Tyr Asn
 Met Ala His Phe Ser Lys Phe Val Arg Pro Gly Tyr Val Arg Ile Asp
 Ala Thr Lys Asn Pro Asn Ala Asn Val Tyr Val Ser Ala Tyr Lys Gly
 Asp Asn Lys Val Val Ile Val Ala Ile Asn Lys Ser Asn Thr Gly Val
 Asn Gln Asn Phe Val Leu Gln Asn Gly Ser Ala Ser Asn Val Ser Arg
 Trp Ile Thr Ser Ser Ser Ser Asn Leu Gln Pro Gly Thr Asn Leu Thr
 Val Ser Gly Asn His Phe Trp Ala His Leu Pro Ala Gln Ser Val Thr
 Thr Phe Val Val Asn Arg
 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420

<210> 239

<211> 1281

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample

<400> 239

atgaatcggt	tcttgatttc	acgttataag	aaagccataa	gtgcatgttt	ggcccttgct	60
cttgcggtgt	ctctcatggc	ggcacctggc	gatgttgccg	cagccagcga	cgccgttata	120
aatgtatcgt	cggagaaaca	agtgatacgc	ggtttcggag	gcatcaacca	cccggcatgg	180
atcggagatt	tgacggcagc	acagagagaa	accgcatttg	ggaacggggc	aatcagttta	240
ggcttctcga	tattaagaat	ctacgtgcat	gaagaccgaa	atcagtgagg	ccgtgaactg	300
gatacggcca	aacgagcgat	tgcccttgga	gctatcgtat	tcgcttcgcc	atggaatccg	360
cccgcggaca	tggtcgagac	cttcaaccgc	aacggcgata	cgtcggcaaa	gcgacttcgt	420

tacgacaagt	ataccgccta	tgcccagcat	cttaacgatt	tcgtaaccta	catgagaaac	480
aatggcgtga	atctctacgc	gatttccgtc	cagaacgagc	ccgattatgc	gcatgactgg	540
acgtggtgga	ctccgcagga	aatgcttcgc	tttatgaaag	aaaatgccgg	atcgatcaac	600
agcagagtga	tcgcaccgga	atcgttccaa	tatctgaaaa	atatgtcggg	cccgattcta	660
aatgatcccc	aggcgcttgc	caatatggat	attcttggcg	ctcatctgta	cggtacccaa	720
gtagcaatt	tcgcttatcc	actattcaaa	caaaaaggag	cgggaaaaga	cctctggatg	780
accgaggtgt	attaccgaa	cagcgacaac	aactcggcgg	atcgctggcc	cgaagccctg	840
gatgtgtctt	accatatcca	caatgcatg	gtagaggag	attttcaagc	ttatgtatgg	900
tggtatatcc	gcagatccta	tggtccaatg	aaagaggacg	gcacgatcag	caaacgcggc	960
tacaatatgg	ctcatttctc	caagtttgtc	cgtcccggct	atgtcagggt	ggatgcttcg	1020
aaaaatccag	aaacgaacgt	ttacgtatcc	gcataataag	gcgacaacaa	aatcgttatc	1080
gttgccataa	accggaacaa	ctccggggtc	aatcagaact	ttgtccttca	gaatggatcc	1140
gtttcgcagg	tatcaagggt	gatcacgagc	agcagcagca	atctccagcc	aggaacgtct	1200
ctcaatgtaa	cagggagcaa	tttctgggct	catcttcccg	cgcaaagcgt	tacgactttt	1260
gtgggtgaac	tcggaaggta	a				1281

<210> 240
 <211> 426
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(30)

<400> 240

Met	Asn	Arg	Phe	Leu	Ile	Ser	Arg	Tyr	Lys	Ala	Ile	Ser	Ala	Cys	
1				5					10				15		
Leu	Ala	Leu	Val	Leu	Ala	Leu	Ser	Leu	Met	Ala	Ala	Pro	Gly	Asp	Val
			20					25					30		
Ala	Ala	Ala	Ser	Asp	Ala	Val	Ile	Asn	Val	Ser	Ser	Glu	Lys	Gln	Val
			35				40					45			
Ile	Arg	Gly	Phe	Gly	Gly	Ile	Asn	His	Pro	Ala	Trp	Ile	Gly	Asp	Leu
	50					55					60				
Thr	Ala	Ala	Gln	Arg	Glu	Thr	Ala	Phe	Gly	Asn	Gly	Pro	Asn	Gln	Leu
65					70				75					80	
Gly	Phe	Ser	Ile	Leu	Arg	Ile	Tyr	Val	His	Glu	Asp	Arg	Asn	Gln	Trp
			85					90					95		
His	Arg	Glu	Leu	Asp	Thr	Ala	Lys	Arg	Ala	Ile	Ala	Leu	Gly	Ala	Ile
			100					105					110		
Val	Phe	Ala	Ser	Pro	Trp	Asn	Pro	Pro	Ala	Asp	Met	Val	Glu	Thr	Phe
			115				120					125			
Asn	Arg	Asn	Gly	Asp	Thr	Ser	Ala	Lys	Arg	Leu	Arg	Tyr	Asp	Lys	Tyr
			130			135					140				
Thr	Ala	Tyr	Ala	Gln	His	Leu	Asn	Asp	Phe	Val	Thr	Tyr	Met	Arg	Asn
145					150					155				160	
Asn	Gly	Val	Asn	Leu	Tyr	Ala	Ile	Ser	Val	Gln	Asn	Glu	Pro	Asp	Tyr
			165					170					175		
Ala	His	Asp	Trp	Thr	Trp	Trp	Thr	Pro	Gln	Glu	Met	Leu	Arg	Phe	Met
			180					185					190		
Lys	Glu	Asn	Ala	Gly	Ser	Ile	Asn	Ser	Arg	Val	Ile	Ala	Pro	Glu	Ser
		195					200					205			
Phe	Gln	Tyr	Leu	Lys	Asn	Met	Ser	Asp	Pro	Ile	Leu	Asn	Asp	Pro	Gln
			210			215					220				
Ala	Leu	Ala	Asn	Met	Asp	Ile	Leu	Gly	Ala	His	Leu	Tyr	Gly	Thr	Gln
225					230					235					240
Val	Ser	Asn	Phe	Ala	Tyr	Pro	Leu	Phe	Lys	Gln	Lys	Gly	Ala	Gly	Lys
			245					250					255		
Asp	Leu	Trp	Met	Thr	Glu	Val	Tyr	Tyr	Pro	Asn	Ser	Asp	Asn	Asn	Ser
			260					265					270		
Ala	Asp	Arg	Trp	Pro	Glu	Ala	Leu	Asp	Val	Ser	Tyr	His	Ile	His	Asn
		275					280					285			
Ala	Met	Val	Glu	Gly	Asp	Phe	Gln	Ala	Tyr	Val	Trp	Trp	Tyr	Ile	Arg
		290				295					300				
Arg	Ser	Tyr	Gly	Pro	Met	Lys	Glu	Asp	Gly	Thr	Ile	Ser	Lys	Arg	Gly
305					310					315					320
Tyr	Asn	Met	Ala	His	Phe	Ser	Lys	Phe	Val	Arg	Pro	Gly	Tyr	Val	Arg

Val Asp Ala Ser 325 Lys Asn Pro Glu Thr 330 Asn Val Tyr Val Ser 335 Ala Tyr
 Lys Gly Asp Asn 340 Lys Ile Val Ile Val 345 Ala Ile Asn Arg Asn 350 Asn Ser
 Gly Val 355 Asn Gln Asn Phe Val 360 Leu Gln Asn Gly Ser Val 365 Ser Gln Val
 Ser Arg Trp Ile Thr Ser 375 Ser Ser Ser Asn Leu 380 Gln Pro Gly Thr Ser
 385 Leu Asn Val Thr Gly 390 Ser Asn Phe Trp Ala His 395 Leu Pro Ala Gln Ser
 Val Thr Thr Phe 405 Val Gly Glu Leu Gly Arg 410 415
 420 425

<210> 241
 <211> 1695
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 241
 gtgaagatat tgaatttaa gatgaattta aaaaaatcgg ttcattgttct gttggcctgt 60
 ttaacagccc tgcctctcat gttaacgccg acacacgtat cagcagcaag tgatgccaac 120
 attaatttgg cctccgaaaa acagcttatc aaggggtttg gaggtattaa ccaccagcc 180
 tggattggcg acttgacggc agctcagcgt gaaacagcct ttggcaacgg agcgaaccag 240
 cttggttttt ccatactaag aatctatgtc gatgaaaatc caaacaactg gtacagggag 300
 gtggctactg ccaaagagc catagagcaa ggtgccatcg tattcgcttc tccctggaat 360
 ccaccaagtg acatggtcga aaccttcaat cggaacgggg atacgaacgc caaacgattg 420
 agatacgaca aatatgctgc gtacgcgcag catctgaacg actttgtcag ttatatgaaa 480
 aataacggtg tggatctgta tgccatttcg gtacaaaatg agccggatta tgcccatgaa 540
 tggacctggg ggactcgcga ggagatcctt cgtttcatga aggagaatgc gggatccatt 600
 cagaatacca aagtcatggc acctgaatcg ttccagtatt tgaaaaacat gtctgaccgg 660
 attctgaatg atcctcaggc actcgccaat atggacattc tgggagctca tacgtacggg 720
 acacagttca aagatttcgc ataccgctc ttttaagcaa agggagccgg caaagaactg 780
 tggatgacag aagtgtatta cccgaacagc gataacaact cgtcggaccg ttggcctgag 840
 gcattggacg tatcttacca tatgcataat gccatgggtg aaggagattt tcaggcttac 900
 gtatgggtgg atattcggag acagtacggg ccgatgaatg agaacgggac tattagcaaa 960
 cgtgggttaca atatggcca tttctccaaa ttgtgctgac caggctatta ccgtgtcgat 1020
 gcaacaaaaa atccggatac caataccttc gtctcagcct ataaagggtg taataaggca 1080
 gtcattgtgg cgattaatcg cggcacctcg gctgtaagcc aaaaattcgt tcttcagaat 1140
 ggtaacgctt ctactgtatc ctcttggtt acggatagca gccgaaacct ggcaagcgga 1200
 gcgcccatta cgatgtcagg tggagccttt acagcacaac tgccagccca aagcgtaaca 1260
 acgtttgtag ccaacttatc tgggtgtagt gtcactccag gcagcggaaac cacgtacgag 1320
 gcggaaacgg gcaactacat taccgatgcc gtgatcgaga ctctctaccc gggatacact 1380
 gggaccggat acgtgaaatt taatgcgtat actgggttcgg ccattcaatg gaatgccatc 1440
 aataacacga taacaggtag caaaaatgtg aaatttcgtt acgcccagga aagcggaaac 1500
 cgtaactctg caactttcgt taacggaaact aaagtcata gcaacgaacc tttcccggca 1560
 acaggcagct ggtcgacctg gagtgaaaaa actattcagg tccccatgaa cgcggaacc 1620
 aatacgatta aagtgtcac aaccggtaca gaagggcca atattgataa catcaatgtc 1680
 actgcagtc aataa 1695

<210> 242
 <211> 564
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(28)

<400> 242
 Val Lys Ile Leu 5 Phe Lys Met Asn Leu Lys Lys Ser Val His Val
 1 10 15
 Leu Leu Ala Cys Leu Thr Ala Leu Pro Leu Met Leu Thr Pro Thr His
 20 25 30

Val Ser Ala Ala Ser Asp Ala Asn Ile Asn Leu Ser Ser Glu Lys Gln
 35 40 45
 Leu Ile Lys Gly Phe Gly Gly Ile Asn His Pro Ala Trp Ile Gly Asp
 50 55 60
 Leu Thr Ala Ala Gln Arg Glu Thr Ala Phe Gly Asn Gly Ala Asn Gln
 65 70 75 80
 Leu Gly Phe Ser Ile Leu Arg Ile Tyr Val Asp Glu Asn Pro Asn Asn
 85 90 95
 Trp Tyr Arg Glu Val Ala Thr Ala Lys Arg Ala Ile Glu Gln Gly Ala
 100 105 110
 Ile Val Phe Ala Ser Pro Trp Asn Pro Pro Ser Asp Met Val Glu Thr
 115 120 125
 Phe Asn Arg Asn Gly Asp Thr Asn Ala Lys Arg Leu Arg Tyr Asp Lys
 130 135 140
 Tyr Ala Ala Tyr Ala Gln His Leu Asn Asp Phe Val Ser Tyr Met Lys
 145 150 155 160
 Asn Asn Gly Val Asp Leu Tyr Ala Ile Ser Val Gln Asn Glu Pro Asp
 165 170 175
 Tyr Ala His Glu Trp Thr Trp Trp Thr Pro Gln Glu Ile Leu Arg Phe
 180 185 190
 Met Lys Glu Asn Ala Gly Ser Ile Gln Asn Thr Lys Val Met Ala Pro
 195 200 205
 Glu Ser Phe Gln Tyr Leu Lys Asn Met Ser Asp Pro Ile Leu Asn Asp
 210 215 220
 Pro Gln Ala Leu Ala Asn Met Asp Ile Leu Gly Ala His Thr Tyr Gly
 225 230 235 240
 Thr Gln Phe Lys Asp Phe Ala Tyr Pro Leu Phe Lys Gln Lys Gly Ala
 245 250 255
 Gly Lys Glu Leu Trp Met Thr Glu Val Tyr Tyr Pro Asn Ser Asp Asn
 260 265 270
 Asn Ser Ser Asp Arg Trp Pro Glu Ala Leu Asp Val Ser Tyr His Met
 275 280 285
 His Asn Ala Met Val Glu Gly Asp Phe Gln Ala Tyr Val Trp Trp Tyr
 290 295 300
 Ile Arg Arg Gln Tyr Gly Pro Met Asn Glu Asn Gly Thr Ile Ser Lys
 305 310 315 320
 Arg Gly Tyr Asn Met Ala His Phe Ser Lys Phe Val Arg Pro Gly Tyr
 325 330 335
 Tyr Arg Val Asp Ala Thr Lys Asn Pro Asp Thr Asn Thr Phe Val Ser
 340 345 350
 Ala Tyr Lys Gly Asp Asn Lys Ala Val Ile Val Ala Ile Asn Arg Gly
 355 360 365
 Thr Ser Ala Val Ser Gln Lys Phe Val Leu Gln Asn Gly Asn Ala Ser
 370 375 380
 Thr Val Ser Ser Trp Val Thr Asp Ser Ser Arg Asn Leu Ala Ser Gly
 385 390 395 400
 Ala Pro Ile Thr Met Ser Gly Gly Ala Phe Thr Ala Gln Leu Pro Ala
 405 410 415
 Gln Ser Val Thr Thr Phe Val Ala Asn Ile Thr Gly Gly Ser Val Thr
 420 425 430
 Pro Gly Ser Gly Thr Thr Tyr Glu Ala Glu Thr Gly Thr Thr Leu Thr
 435 440 445
 Asp Ala Val Ile Glu Thr Leu Tyr Pro Gly Tyr Thr Gly Thr Gly Tyr
 450 455 460
 Val Asn Phe Asn Ala Tyr Thr Gly Ser Ala Ile Gln Trp Asn Ala Ile
 465 470 475 480
 Asn Asn Thr Ile Thr Gly Thr Lys Asn Val Lys Phe Arg Tyr Ala Gln
 485 490 495
 Glu Ser Gly Thr Arg Asn Leu Asp Ile Phe Val Asn Gly Thr Lys Val
 500 505 510
 Ile Ser Asn Glu Pro Phe Pro Ala Thr Gly Ser Trp Ser Thr Trp Ser
 515 520 525
 Glu Lys Thr Ile Gln Val Pro Met Asn Ala Gly Thr Asn Thr Ile Lys
 530 535 540
 Val Val Thr Thr Gly Thr Glu Gly Pro Asn Ile Asp Asn Ile Asn Val
 545 550 555 560
 Thr Ala Val Gln

<210> 243
 <211> 1272
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 243
 atgatttcaa gcgtaaaaa accaatttgt gtattattgg tatgtttcac tatgctgtca 60
 gtcatgttag ccgggccagg tgctactgaa gttttagcag caagtgatgt aacaattaat 120
 ttatctgcag aaaaacaagt gatccgcggt tttggaggca tgaaccaccc ggcttggatt 180
 ggagatttga cagcagctca aagagaaacc gcttttggca atggacagaa tcagttaggt 240
 ttttcaatct taagaattca tgtggatgaa aatagaaata attggtacag agaagtggag 300
 actgcaaaga gtgcatcaa acatggagca atcgtttttg cttctccctg gaatcctcca 360
 agcgatatgg ttgagacttt caatcgtaat ggtgacacat cagctaaacg gctaagatac 420
 gataagtacg ccgcatacgc gcagcatctt aacgattttg ttacctacat gaagaataat 480
 ggcgtgaatc tttatgcgat ttctgttcaa aacgagcctg attatgcgca cgaatggacg 540
 tgggtggactc cgcaagaaat acttcgtttc atgagagaaa atgccgggtc cattaatgca 600
 cgtgtcattg caccagaatc ttttcagtac tttaaaata tatcggaccc cattttgaac 660
 gatccacagg cgcttaggaa tatggatatt ctcggaactc acctgtacgg tactcaggtc 720
 agtcagtttc atttctcgaa attcaaacaa aaaggagcag ggaaagagct atggatgacg 780
 gaagtatact atccaaacag tgacaacaat tcagcggatc gctggcccga ggcattaggc 840
 gtttcagagc atattcacca ttcaatgggtg gagggagatt ttcaatctta tgtttgggtg 900
 tacatccgca gatcttacgg tcctatgaaa gaggacggta cgatcagcaa acgcggttac 960
 aatatggctc atttctcgaa gtttgtgcgt cccggctatg taagggtaga tgcaacgaaa 1020
 aatcctaata cgaacgttta cgtgtcagcc tataaagggt acaacaaggt cgttattggt 1080
 gccattaaca aaagcaatac aggggtcaac caaaactttg tgttgcagaa tggatctgct 1140
 tctcaggtat ctaggtggat aacaagcgga agcagcaatc ttcaacctgg aacgaatctc 1200
 aatgtaacgg gcaatcattt ttgggcccac cttccagctc aaagcgtgac aacatttgtc 1260
 gcaaatcgtt aa 1272

<210> 244
 <211> 423
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(33)

<400> 244
 Met Ile Ser Ser Val Lys Lys Pro Ile Cys Val Leu Leu Val Cys Phe
 1 5 10 15
 Thr Met Leu Ser Val Met Leu Ala Gly Pro Gly Ala Thr Glu Val Leu
 20 25 30
 Ala Ala Ser Asp Val Thr Ile Asn Leu Ser Ala Glu Lys Gln Val Ile
 35 40 45
 Arg Gly Phe Gly Gly Met Asn His Pro Ala Trp Ile Gly Asp Leu Thr
 50 55 60
 Ala Ala Gln Arg Glu Thr Ala Phe Gly Asn Gly Gln Asn Gln Leu Gly
 65 70 75 80
 Phe Ser Ile Leu Arg Ile His Val Asp Glu Asn Arg Asn Asn Trp Tyr
 85 90 95
 Arg Glu Val Glu Thr Ala Lys Ser Ala Ile Lys His Gly Ala Ile Val
 100 105 110
 Phe Ala Ser Pro Trp Asn Pro Pro Ser Asp Met Val Glu Thr Phe Asn
 115 120 125
 Arg Asn Gly Asp Thr Ser Ala Lys Arg Leu Arg Tyr Asp Lys Tyr Ala
 130 135 140
 Ala Tyr Ala Gln His Leu Asn Asp Phe Val Thr Tyr Met Lys Asn Asn
 145 150 155 160
 Gly Val Asn Leu Tyr Ala Ile Ser Val Gln Asn Glu Pro Asp Tyr Ala
 165 170 175
 His Glu Trp Thr Trp Thr Pro Gln Ile Leu Arg Phe Met Arg
 180 185 190
 Glu Asn Ala Gly Ser Ile Asn Ala Arg Val Ile Ala Pro Glu Ser Phe
 Page 177

195 200 205
 Gln Tyr Phe Lys Asn Ile Ser Asp Pro Ile Leu Asn Asp Pro Gln Ala
 210 215 220
 Leu Arg Asn Met Asp Ile Leu Gly Thr His Leu Tyr Gly Thr Gln Val
 225 230 235 240
 Ser Gln Phe Pro Tyr Pro Leu Phe Lys Gln Lys Gly Ala Gly Lys Glu
 245 250 255
 Leu Trp Met Thr Glu Val Tyr Tyr Pro Asn Ser Asp Asn Asn Ser Ala
 260 265 270
 Asp Arg Trp Pro Glu Ala Leu Gly Val Ser Glu His Ile His His Ser
 275 280 285
 Met Val Glu Gly Asp Phe Gln Ser Tyr Val Trp Trp Tyr Ile Arg Arg
 290 295 300
 Ser Tyr Gly Pro Met Lys Glu Asp Gly Thr Ile Ser Lys Arg Gly Tyr
 305 310 315 320
 Asn Met Ala His Phe Ser Lys Phe Val Arg Pro Gly Tyr Val Arg Val
 325 330 335
 Asp Ala Thr Lys Asn Pro Asn Ala Asn Val Tyr Val Ser Ala Tyr Lys
 340 345 350
 Gly Asp Asn Lys Val Val Ile Val Ala Ile Asn Lys Ser Asn Thr Gly
 355 360 365
 Val Asn Gln Asn Phe Val Leu Gln Asn Gly Ser Ala Ser Gln Val Ser
 370 375 380
 Arg Trp Ile Thr Ser Gly Ser Ser Asn Leu Gln Pro Gly Thr Asn Leu
 385 390 395 400
 Asn Val Thr Gly Asn His Phe Trp Ala His Leu Pro Ala Gln Ser Val
 405 410 415
 Thr Thr Phe Val Ala Asn Arg
 420

<210> 245
 <211> 1263
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 245
 atgtcaatga tcaaaaaacc aatctgcact ttattgatct gcttcaccat gctgtctgtc 60
 atgttcatcg gacctggcgt gactgaggtt tcagcagcag atgccaatat taatatcaat 120
 gcggaaagac aagtgattcg cggctttggc ggaatgaacc atccggcttg gattggtgat 180
 ttgaccgcac ctcaaagggg aaccgccttt ggcaatgggc agaatcaatt aggattttcc 240
 attctaagaa tttttgtaga tgagaaccga aataattggc acagagaggt cgctactgcc 300
 aaaagagcaa tagagcatgg agctttgggtg atcgcttcac catggaatcc tccaagcaat 360
 atggttagaga ccttcaaccg gaatggtaca tctgcaaagc ggctcagata caaccaatac 420
 gccgcatatg ctccagcatct gaacgatttt gtgacgtata tgaaaaataa tggcgtcaat 480
 ctctatgcta tatctgtaca aaatgagccc gattatgcac acgaatggac atgggtggact 540
 cctcaggaaa tcctgcgttt catgagagaa aatgctggct ccattaatgc ccgcgtgatc 600
 gcaccagaat cctttcaata ctttaaaaat atatcagatc ctatcctaaa cgatccgcag 660
 gcgcttggaac acatggacat tctcggagcc catttgtagc gaaccctaat cagccagctt 720
 ccgtatcctc ttttcaaaca aaagggaggg ggaaaggagc tttggatgac agaggtctac 780
 taccggaata gcgataacaa ttcagcggac cgctggcctg aagcattagg ggtttcagag 840
 catattcacc attcgatggg agaaggggac tttcaggcat atgtttgggt gtacattcgc 900
 agatcctacg gccctatgaa agaagacggt ctaatcagca aacgtgggta caacatggcg 960
 cattttctcca agttttgtacg cccaggatac atcagaattg atgcaacgaa aagtcctgaa 1020
 ccgaatgttt tcgtagtcagc ctataaagga aacaatcaag tcgtcattgt cgcgattaac 1080
 aaaaacaata caggagtcaa tcagcacttt gtgatgcaaa acggaactgc ttcacaagcg 1140
 tcaagatgga tcacaagtag caacagcaac cttcagcctg gtacagactt aaatatatca 1200
 ggtaatcaat tttgggctca tctcccggct caaagtgtga caacatttgt ggtcaaacgc 1260
 tag 1263

<210> 246
 <211> 401
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample
 Page 178

<221> SIGNAL
 <222> (1)...(32)

<400> 246

```

Met Ser Met Ile Lys Lys Pro Ile Cys Thr Leu Leu Ile Cys Phe Thr
1      5      10      15
Met Leu Ser Val Met Phe Ile Gly Pro Gly Val Thr Glu Val Ser Ala
20      25      30
Ala Asp Ala Asn Ile Asn Ile Asn Ala Glu Arg Gln Val Ile Arg Gly
35      40      45
Phe Gly Gly Met Asn His Pro Ala Trp Ile Gly Asp Leu Thr Ala Pro
50      55      60
Gln Arg Glu Thr Ala Phe Gly Asn Gly Gln Asn Gln Leu Gly Phe Ser
65      70      75      80
Ile Leu Arg Ile Phe Val Asp Glu Asn Arg Asn Asn Trp His Arg Glu
85      90      95
Val Ala Thr Ala Lys Arg Ala Ile Glu His Gly Ala Leu Val Ile Ala
100     105     110
Ser Pro Trp Asn Pro Pro Ser Asn Met Val Glu Thr Phe Asn Arg Asn
115     120     125
Gly Thr Ser Ala Lys Arg Leu Arg Tyr Asn Gln Tyr Ala Ala Tyr Ala
130     135     140
Gln His Leu Asn Asp Phe Val Thr Tyr Met Lys Asn Asn Gly Val Asn
145     150     155     160
Leu Tyr Ala Ile Ser Val Gln Asn Glu Pro Asp Tyr Ala His Glu Trp
165     170     175
Thr Trp Trp Thr Pro Gln Glu Ile Leu Arg Phe Met Arg Glu Asn Ala
180     185     190
Gly Ser Ile Asn Ala Arg Val Ile Ala Pro Glu Ser Phe Gln Tyr Leu
195     200     205
Lys Asn Ile Ser Asp Pro Ile Leu Asn Asp Pro Gln Ala Leu Gly Asn
210     215     220
Met Asp Ile Leu Gly Ala His Leu Tyr Gly Thr Gln Ile Ser Gln Leu
225     230     235     240
Pro Tyr Pro Leu Phe Lys Gln Lys Gly Gly Gly Lys Glu Leu Trp Met
245     250     255
Thr Glu Val Tyr Tyr Pro Asn Ser Asp Asn Asn Ser Ala Asp Arg Trp
260     265     270
Pro Glu Ala Leu Gly Val Ser Glu His Ile His His Ser Met Val Glu
275     280     285
Gly Asp Phe Gln Ala Tyr Val Trp Trp Tyr Ile Arg Arg Ser Tyr Gly
290     295     300
Pro Met Lys Glu Asp Gly Leu Ile Ser Lys Arg Gly Tyr Asn Met Ala
305     310     315     320
His Phe Ser Lys Phe Val Arg Pro Gly Tyr Ile Arg Ile Asp Ala Thr
325     330     335
Lys Ser Pro Glu Pro Asn Val Phe Val Ser Ala Tyr Lys Gly Asn Asn
340     345     350
Gln Val Val Ile Val Ala Ile Asn Lys Asn Asn Thr Gly Val Asn Gln
355     360     365
His Phe Val Met Gln Asn Gly Thr Ala Ser Gln Ala Ser Arg Trp Ile
370     375     380
Thr Ser Ser Asn Ser Asn Leu Gln Pro Gly Thr Asp Leu Asn Ile Ser
385     390     395     400
Gly

```

<210> 247
 <211> 1044
 <212> DNA
 <213> Unknown

<220>

<223> obtained from an environmental sample

<400> 247

```

gtgttttgcca acgatttcct gataggcggtg gcgctcaact cacggcaggt cgccggggaa    60
tccgaggccg gaaaactagc tggcgcgcaa ttttcgtcgg tgacggcgga gaatgagatg    120

```

aagtggcagt	cgctccatcc	ccagcccgac	cgctatcagt	tcggcgcggc	ggactcctac	180
atcgattttg	ccaaaaaaca	caagatggcg	gtgatcggcc	acacgctcgt	gtggcacagc	240
cagacacccg	gctgggtggt	cgagggcaag	gacggcaagc	cggcgacccg	cgaggatctg	300
ctcaagcgca	tgcgcgatca	catccacacc	gtggccggac	gctacaaggg	caaggtgcgc	360
ggctgggacg	tggtcaacga	ggccttgtcc	gacggcggtc	ccgaaatcct	gcgggattct	420
ccgtggcggc	gcatcatcgg	cgatgacttc	atcgaccacg	cgttccgttt	cgcccgtgag	480
gccgatccga	aagccgaact	ctactacaac	gactacggtc	tcgagaacga	aaggaagcgg	540
agcaactgca	tcaagctcgt	caagggcatg	aaacaacgcg	gcgtgccgat	cgacgggggtg	600
ggcaccacgt	cgcattttcca	cttgaaacat	ccctcgctcc	aggaaatcga	aaagaccatc	660
aaggactttt	ccgaactcgg	actcaagggtg	atgatcacccg	agctggatgt	cgatgtgctg	720
ccgtcgcggtg	gcaatttcgg	caacgccgac	atcagccgcc	gcgagcaggg	cggtgacgca	780
ctcaatcctt	acaccggcgg	cttgcccgat	gaggtccaac	aggaacttgc	gaaacgctat	840
gcggacattt	ttgatatcta	tctgcgccac	cgggaaggcgg	tcacccgcgt	aaccttctgg	900
ggactcgatg	acgggcatac	ctggttgaac	ggtttcccga	tcgcggacg	caccaactat	960
ccgctgttgt	tcgaccgcgc	cctcaagccg	aagcccgcgt	tcgaggcggt	catcaaaaaa	1020
gggcttgaac	ccaggaaacg	ttga				1044

<210> 248

<211> 347

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 248

Val	Phe	Ala	Asn	Asp	Phe	Leu	Ile	Gly	Val	Ala	Leu	Asn	Ser	Arg	Gln
1				5					10					15	
Val	Ala	Gly	Glu	Ser	Glu	Ala	Gly	Lys	Leu	Ala	Gly	Ala	Gln	Phe	Ser
			20					25					30		
Ser	Val	Thr	Ala	Glu	Asn	Glu	Met	Lys	Trp	Gln	Ser	Leu	His	Pro	Gln
		35					40					45			
Pro	Asp	Arg	Tyr	Gln	Phe	Gly	Ala	Ala	Asp	Ser	Tyr	Ile	Asp	Phe	Ala
	50					55					60				
Lys	Lys	His	Lys	Met	Ala	Val	Ile	Gly	His	Thr	Leu	Val	Trp	His	Ser
65					70				75					80	
Gln	Thr	Pro	Gly	Trp	Val	Phe	Glu	Gly	Lys	Asp	Gly	Lys	Pro	Ala	Thr
			85						90					95	
Arg	Glu	Asp	Leu	Leu	Lys	Arg	Met	Arg	Asp	His	Ile	His	Thr	Val	Ala
			100					105					110		
Gly	Arg	Tyr	Lys	Gly	Lys	Val	Arg	Gly	Trp	Asp	Val	Val	Asn	Glu	Ala
		115					120					125			
Leu	Ser	Asp	Gly	Gly	Pro	Glu	Ile	Leu	Arg	Asp	Ser	Pro	Trp	Arg	Arg
		130				135					140				
Ile	Ile	Gly	Asp	Asp	Phe	Ile	Asp	His	Ala	Phe	Arg	Phe	Ala	Arg	Glu
145					150				155					160	
Ala	Asp	Pro	Lys	Ala	Glu	Leu	Tyr	Tyr	Asn	Asp	Tyr	Gly	Leu	Glu	Asn
			165						170					175	
Glu	Arg	Lys	Arg	Ser	Asn	Cys	Ile	Lys	Leu	Val	Lys	Gly	Met	Lys	Gln
			180					185					190		
Arg	Gly	Val	Pro	Ile	Asp	Gly	Val	Gly	Thr	Gln	Ser	His	Phe	His	Leu
		195					200					205			
Lys	His	Pro	Ser	Leu	Gln	Glu	Ile	Glu	Lys	Thr	Ile	Lys	Asp	Phe	Ser
		210				215					220				
Glu	Leu	Gly	Leu	Lys	Val	Met	Ile	Thr	Glu	Leu	Asp	Val	Asp	Val	Leu
225					230					235				240	
Pro	Ser	Arg	Gly	Asn	Phe	Gly	Asn	Ala	Asp	Ile	Ser	Arg	Arg	Glu	Gln
			245						250					255	
Gly	Gly	Asp	Ala	Leu	Asn	Pro	Tyr	Thr	Gly	Gly	Leu	Pro	Asp	Glu	Val
			260					265					270		
Gln	Gln	Glu	Leu	Ala	Lys	Arg	Tyr	Ala	Asp	Ile	Phe	Asp	Ile	Tyr	Leu
		275					280					285			
Arg	His	Arg	Lys	Ala	Val	Thr	Arg	Val	Thr	Phe	Trp	Gly	Leu	Asp	Asp
		290				295					300				
Gly	His	Thr	Trp	Leu	Asn	Gly	Phe	Pro	Ile	Arg	Gly	Arg	Thr	Asn	Tyr
305					310					315				320	
Pro	Leu	Leu	Phe	Asp	Arg	Ala	Leu	Lys	Pro	Lys	Pro	Ala	Phe	Glu	Ala
			325						330					335	
Val	Ile	Lys	Lys	Gly	Leu	Glu	Pro	Arg	Lys	Arg					

340

345

<210> 249
 <211> 1439
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 249
 tgatcaatcc agtgaaggat cttcgtgaag atttcatctt tggaatggac gtttcaatgc 60
 tctacgagat agagcggctc ggtggttaagt tcttcgatgg tgggtgtggag aaagatcttt 120
 tccagatact gaaggatcat gagataaact ggatcagatt gagagtgtgg aacgatccaa 180
 gggatgaaaa cggaaatccg ctccggtgggg gaaactgtga ttatctgaaa atgacagaga 240
 tcgcaaaaag ggccaaaaag tacggaatga aggttcttct tgactttcac tacagcgact 300
 ggtgggcaga tcccggcaag cagtacaaac caaaagagtg ggatcacctt catggagaac 360
 ttctggaaaag ggcggtgtat tcctacacga aactcgtgct gaatcatatg agaagaaacg 420
 gtgcaactgcc ggacatggtc caggtgggaa acgaggtgaa caacggcttt ctctggcccg 480
 atggaatgat tgccggaaag gatgcaggag gattcgacgg attcacaaaa cttttgaagg 540
 cggccatcaa agccgtcagg gaagttgatc ccgatatcaa gatagtcatt catttggcag 600
 aaggtggaaa caactcactt ttcaggtggt tcttcgacga gatacacaaga agagacgtgg 660
 attttgatgt gatcgggtga tcgtactatc cgtactggca tgggtaccctg gatgacctga 720
 agaacaacct gtacgacata gcgaaaagat acaacaaaga cgtgctcatc gttgaaacgg 780
 cgtatgcctg gacactcgag gacggggacg gttaccccaa catcttcagt ggtgaagaga 840
 tggagctcac ggggtggttac aaagcaacgg ttcagggaca ggcaacgttc ttgagggatc 900
 tcatagaagt ggtgaacagt gttcctgacg gtcaaggctt tgggatcttc tattgggaag 960
 gagactggat tcctgtgaaa ggagccggct ggaaaaccgg cgaaggaaat ccatgggaga 1020
 atcaggccat gtttgatttc aatggaaatg ctctcccatc cctggatgtt ttcaagctcg 1080
 tgaggacagt cactcctatg gaaataaaaa tcgaagagat tctgcctgtg gagatctcga 1140
 cgaatttggg agagattccg aagtttccgg atgctgtgaa agtgctgttc agcgatgatt 1200
 ccatcagatc cctgaaagtt acatggaatt ttgatccttc tcttgttgaa acacccgggtg 1260
 tctacagagt ggaaggatac gtggaaagta tagaccagaa gatcttcgca accttgactg 1320
 tgaagggaag tagaaactac ctgaagaacc ctgggtttga aacgggtgag ttttctccct 1380
 ggaagggtgt cggtaacgga aaacgcagtg aaggtggtaa aggccgatcc tccgagtaa 1439

<210> 250
 <211> 479
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(33)

<400> 250
 Met Ile Asn Pro Val Lys Asp Leu Arg Glu Asp Phe Ile Phe Gly Met
 1 5 10 15
 Asp Val Ser Met Leu Tyr Glu Ile Glu Arg Leu Gly Gly Lys Phe Phe
 20 25 30
 Asp Gly Gly Val Glu Lys Asp Leu Phe Gln Ile Leu Lys Asp His Glu
 35 40 45
 Ile Asn Trp Ile Arg Leu Arg Val Trp Asn Asp Pro Arg Asp Glu Asn
 50 55 60
 Gly Asn Pro Leu Gly Gly Glu Asn Cys Asp Tyr Leu Lys Met Thr Glu
 65 70 75 80
 Ile Ala Lys Arg Ala Lys Lys Tyr Gly Met Lys Val Leu Leu Asp Phe
 85 90 95
 His Tyr Ser Asp Trp Trp Ala Asp Pro Gly Lys Gln Tyr Lys Pro Lys
 100 105 110
 Glu Trp Asp His Leu His Gly Glu Leu Leu Glu Arg Ala Val Tyr Ser
 115 120 125
 Tyr Thr Lys Leu Val Leu Asn His Met Arg Arg Asn Gly Ala Leu Pro
 130 135 140
 Asp Met Val Gln Val Gly Asn Glu Val Asn Asn Gly Phe Leu Trp Pro
 145 150 155 160
 Asp Gly Met Ile Ala Gly Lys Asp Ala Gly Gly Phe Asp Gly Phe Thr

Lys Leu Leu Lys 165 Ala Ala Ile Lys Ala 170 Val Arg Glu Val Asp 175 Pro Asp
 Ile Lys Ile Val 180 Ile His Leu Ala 185 Glu Gly Gly Asn Asn Ser Leu Phe
 Arg Trp Phe 195 Phe Asp Glu Ile 200 Thr Arg Arg Asp Val 205 Asp Phe Asp Val
 Ile Gly Val Ser Tyr Tyr 215 Pro Tyr Trp His Gly Thr Leu Asp Asp Leu
 225 Lys Asn Asn Leu Tyr Asp Ile Ala Lys Arg Tyr Asn Lys Asp Val Leu
 Ile Val Glu Thr 245 Ala Tyr Ala Trp Thr 250 Leu Glu Asp Gly Asp Gly Tyr
 Pro Asn Ile Phe Ser Gly Glu Glu Met Glu Leu Thr Gly Gly Tyr Lys
 Ala Thr Val Gln Gly Gln Ala Thr Phe Leu Arg Asp Leu Ile Glu Val
 Val Asn Ser Val Pro Asp Gly His Gly Leu Gly Ile Phe Tyr Trp Glu
 305 Gly Asp Trp Ile Pro Val Lys Gly Ala Gly Trp Lys Thr Gly Glu Gly
 Asn Pro Trp Glu Asn Gln Ala Met Phe Asp Phe Asn Gly Asn Ala Leu
 Pro Ser Leu Asp Val Phe Lys Leu Val Arg Thr Val Thr Pro Met Glu
 Ile Lys Ile Glu Glu Ile Leu Pro Val Glu Ile Ser Thr Asn Leu Gly
 Glu Ile Pro Lys Phe Pro Asp Ala Val Lys Val Leu Phe Ser Asp Asp
 385 Ser Ile Arg Ser Leu Lys Val Thr Trp Asn Phe Asp Pro Ser Leu Val
 Glu Thr Pro Gly Val Tyr Arg Val Glu Gly Tyr Val Glu Ser Ile Asp
 Gln Lys Ile Phe Ala Thr Leu Thr Val Lys Gly Ser Arg Asn Tyr Leu
 Lys Asn Pro Gly Phe Glu Thr Gly Glu Phe Ser Pro Trp Lys Val Phe
 450 Gly Asn Gly Lys Arg Ser Glu Gly Gly Lys Gly Arg Ser Ser Glu
 465 470 475

<210> 251
 <211> 555
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 251						60
atggctacgg	attattggca	atattggacg	gatggcggcg	gaacggtgaa	tgcgggttaac	120
gggtccgggg	gcaattacag	cgtaacttgg	caaaatagcg	gggacttcgt	ggtcggcaaa	180
ggctggagcg	tagggtcgcc	aaatcggacg	atcaattaca	atgccggcat	ctgggaacct	240
tcggggaacg	ggtacttgac	cctttacgga	tgactagaa	actcgctgat	cgagtattac	300
gtgttcgaca	gttgggggac	gtaccggcca	acaggtactc	acaaaggaac	ggtgaacagc	360
gacggaggca	cctacgatat	ttatacgacc	atgcgctata	atgcgccttc	cattgatggc	420
acgcagacgt	tccaacagtt	ctggagcggtg	cggaatcga	aacgaccaac	cggcagcaac	480
gtctccatca	ccttcagcaa	tcacgtgaat	gcctggagaa	gcaagggcat	gaacctgggc	540
agcagctggg	cgtaccaggt	cttggcgacg	gaaggctatc	agagcagcgg	aagatccaac	555
gtcacggtgt	ggtaa					

<210> 252
 <211> 184
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<400> 252

Met Ala Thr Asp Tyr Trp Gln Tyr Trp Thr Asp Gly Gly Gly Thr Val
 1 5 10 15
 Asn Ala Val Asn Gly Ser Gly Gly Asn Tyr Ser Val Thr Trp Gln Asn
 20 25 30
 Ser Gly Asp Phe Val Val Gly Lys Gly Trp Ser Val Gly Ser Pro Asn
 35 40 45
 Arg Thr Ile Asn Tyr Asn Ala Gly Ile Trp Glu Pro Ser Gly Asn Gly
 50 55 60
 Tyr Leu Thr Leu Tyr Gly Trp Thr Arg Asn Ser Leu Ile Glu Tyr Tyr
 65 70 75 80
 Val Val Asp Ser Trp Gly Thr Tyr Arg Pro Thr Gly Thr His Lys Gly
 85 90 95
 Thr Val Asn Ser Asp Gly Gly Thr Tyr Asp Ile Tyr Thr Thr Met Arg
 100 105 110
 Tyr Asn Ala Pro Ser Ile Asp Gly Thr Gln Thr Phe Gln Gln Phe Trp
 115 120 125
 Ser Val Arg Gln Ser Lys Arg Pro Thr Gly Ser Asn Val Ser Ile Thr
 130 135 140
 Phe Ser Asn His Val Asn Ala Trp Arg Ser Lys Gly Met Asn Leu Gly
 145 150 155 160
 Ser Ser Trp Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln Ser Ser
 165 170 175
 Gly Arg Ser Asn Val Thr Val Trp
 180

<210> 253
 <211> 1047
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<400> 253
 atgattgtta gcttcaagag cctgaaggca ctgcggtgct tcggcggtgct cggcatcacc 60
 gccgcgcacg cgaaacctg catcacgtcg agccagacgg gcaccaacaa cggcaattac 120
 ttttcgttct ggaaagacag tccgggcacg gtgaacttct gcatgtatgc gaatggccgc 180
 tatacctcca actggagcgg catcaacaac tgggtgggcg gcaagggctg ggctaccggc 240
 tccagccaca cgatcagcta ctccggcacg ttcaattcgc cgggcaacgg ttacctggcc 300
 ctgtatggct ggaccaccaa tccattgggtc gagtactaca tcgtcgacag ctgggggtacc 360
 taccgtccgc cgggcggcca gggtttcatg ggcacggtag ttagcgacgg gggcacgtac 420
 gacgtgtacc ggacgcaacg cgtgaaccag ccatccatca tcggcaacgc cacgttctac 480
 cagtactgga gctgcgcgga gtcgaagcgc gtgggcggca ccatcaccat cgccaacat 540
 ttcaacgcct gggccacgct gggcatgaac ctgggccagc acaactacca ggtcatggcc 600
 accgaggggt accagagcag cggcagctcc gacatcaccg tgaccgaagg tggcggcagy 660
 tcctcgtcgt cctcgggcgg cggcagcacc agcagcagtg gtggcggcgg caacaagagc 720
 ttcacgggtgc gtgcgcgcgg cacggccgga ggcgagaaca tccagctgca ggtgaacaac 780
 cagacggctc cgagctggaa cctcaccacc agcatgcaga actacaccgc ctcgaccagc 840
 ctgagcggcg gcatcacctg gctctacacc aacgacggcg gcagccgcga cgtgcagggt 900
 gactacatca tcgtgaacgg ccagacccgc cagtccgaag cgagagcta caacaccggg 960
 ttgtatgcga atggacgctg cggcgggtggc tcgaacagcg agtggatgca ttgcaacggc 1020
 gcgatcggct acggcaatac gccctga 1047

<210> 254
 <211> 347
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(24)

<400> 254
 Met Ile Val Ser Phe Lys Ser Leu Lys Ala Leu Ala Cys Leu Gly Val
 1 5 10 15
 Leu Gly Ile Thr Ala Ala His Ala Gln Thr Cys Ile Thr Ser Ser Gln
 20 25 30

Thr Gly Thr Asn Asn Gly Asn Tyr Phe Ser Phe Trp Lys Asp Ser Pro
 35 40 45
 Gly Thr Val Asn Phe Cys Met Tyr Ala Asn Gly Arg Tyr Thr Ser Asn
 50 55 60
 Trp Ser Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Ala Thr Gly
 65 70 75 80
 Ser Ser His Thr Ile Ser Tyr Ser Gly Thr Phe Asn Ser Pro Gly Asn
 85 90 95
 Gly Tyr Leu Ala Leu Tyr Gly Trp Thr Thr Asn Pro Leu Val Glu Tyr
 100 105 110
 Tyr Ile Val Asp Ser Trp Gly Thr Tyr Arg Pro Pro Gly Gly Gln Gly
 115 120 125
 Phe Met Gly Thr Val Val Ser Asp Gly Gly Thr Tyr Asp Val Tyr Arg
 130 135 140
 Thr Gln Arg Val Asn Gln Pro Ser Ile Ile Gly Asn Ala Thr Phe Tyr
 145 150 155 160
 Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg Val Gly Gly Thr Ile Thr
 165 170 175
 Ile Ala Asn His Phe Asn Ala Trp Ala Thr Leu Gly Met Asn Leu Gly
 180 185 190
 Gln His Asn Tyr Gln Val Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly
 195 200 205
 Ser Ser Asp Ile Thr Val Thr Glu Gly Gly Gly Ser Ser Ser Ser Ser
 210 215 220
 Gly Gly Gly Ser Thr Ser Ser Gly Gly Gly Gly Asn Lys Ser Phe
 225 230 235 240
 Thr Val Arg Ala Arg Gly Thr Ala Gly Gly Glu Asn Ile Gln Leu Gln
 245 250 255
 Val Asn Asn Gln Thr Val Ala Ser Trp Asn Leu Thr Thr Ser Met Gln
 260 265 270
 Asn Tyr Thr Ala Ser Thr Ser Leu Ser Gly Gly Ile Thr Val Leu Tyr
 275 280 285
 Thr Asn Asp Gly Gly Ser Arg Asp Val Gln Val Asp Tyr Ile Ile Val
 290 295 300
 Asn Gly Gln Thr Arg Gln Ser Glu Ala Gln Ser Tyr Asn Thr Gly Leu
 305 310 315 320
 Tyr Ala Asn Gly Arg Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His
 325 330 335
 Cys Asn Gly Ala Ile Gly Tyr Gly Asn Thr Pro
 340 345

<210> 255

<211> 1137

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 255

ttgatctttt	ccgtcagtgg	ttccgcgtct	cggcggcgcc	ctggcatcca	caaggggggat	60
tccatgattt	tccgttctaaa	gtcgatcacg	ggcaggcgcg	ccgtcgcggc	gctggcctgc	120
cttgccggcc	tctacatggc	gccggcgaat	gcgcaaacct	gcatcacgtc	gagccagacg	180
ggcaccaaca	acggcaacta	cttttcgttc	tggaagagaca	gcccgggcac	ggtgaacttc	240
tgcattgtact	ccggcggccg	ctacacgtcc	aactggagcg	gcatcaacaa	ctgggtgggc	300
ggcaagggct	ggcagacggg	ctcgtcccgc	accgtctcct	actccggcag	cttcaattcg	360
ccgggtaacg	gctacctgac	gctctacggc	tggaaccacca	atccggtcat	cgagtactac	420
atcgtcgcaca	actggggcag	ctatcgtccg	ccgggtggcc	agggtttcat	gggcacgggtg	480
aacaccgacg	gcggcagcta	cgacatctat	cgcacgcaac	gggtcaacca	gccgtcgatc	540
atcggcaccg	cgacgtttcta	ccagtactgg	agcgtgcggc	agtcgaagcg	caccggcgcc	600
accatcacca	cggccaacca	cttcaatgcc	tgggcccagcc	tcggcatgaa	cctgggacag	660
cacaactacc	aggtgatggc	caccgagggc	taccagagca	gcggcagctc	cgacatcacg	720
gtgtgggaag	gcagcgccgg	cggcggaagc	agcaatggcg	gcagcagcaa	cggcggcagc	780
agcaatgggtg	gcagcgccgg	cacgaagagc	ttcacgggtc	gcgcgcgcgg	cactgcgggc	840
ggcagagtcca	tcacgctgcg	ggtcaacaac	cagaacgtgc	agacctggac	gctgggtacc	900
agcatgcaga	actacacggc	ctcgacctcg	ctgagcggcg	gcatcacggt	ggcgttcacc	960
aacgacggcg	gcagcgcgca	cgtgcaggtg	gactacatca	tcgtgaatgg	ccagaccgcc	1020
cagtccgaac	agcagagcta	caactactgg	ctctacgccca	atggaagctg	tggtggcggt	1080
tcgaacagcg	agtggatgca	ttgcaacggc	gccatcggtc	acggcaatac	gccctga	1137

<210> 256
 <211> 378
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample

<221> SIGNAL
 <222> (1)...(51)

<400> 256
 Leu Ile Phe Ser Val Ser Gly Ser Ala Ser Arg Arg Arg Pro Gly Ile
 1 5 10 15
 His Lys Gly Asp Ser Met Ile Phe Gly Leu Lys Ser Ile Thr Gly Arg
 20 25 30
 Arg Ala Val Ala Ala Leu Ala Cys Leu Ala Gly Leu Tyr Met Ala Pro
 35 40 45
 Ala Asn Ala Gln Thr Cys Ile Thr Ser Ser Gln Thr Gly Thr Asn Asn
 50 55 60
 Gly Asn Tyr Phe Ser Phe Trp Lys Asp Ser Pro Gly Thr Val Asn Phe
 65 70 75 80
 Cys Met Tyr Ser Gly Gly Arg Tyr Thr Ser Asn Trp Ser Gly Ile Asn
 85 90 95
 Asn Trp Val Gly Gly Lys Gly Trp Gln Thr Gly Ser Ser Arg Thr Val
 100 105 110
 Ser Tyr Ser Gly Ser Phe Asn Ser Pro Gly Asn Gly Tyr Leu Thr Leu
 115 120 125
 Tyr Gly Trp Thr Thr Asn Pro Leu Ile Glu Tyr Tyr Ile Val Asp Asn
 130 135 140
 Trp Gly Ser Tyr Arg Pro Pro Gly Gly Gln Gly Phe Met Gly Thr Val
 145 150 155 160
 Asn Thr Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Thr Gln Arg Val Asn
 165 170 175
 Gln Pro Ser Ile Ile Gly Thr Ala Thr Phe Tyr Gln Tyr Trp Ser Val
 180 185 190
 Arg Gln Ser Lys Arg Thr Gly Gly Thr Ile Thr Thr Ala Asn His Phe
 195 200 205
 Asn Ala Trp Ala Ser Leu Gly Met Asn Leu Gly Gln His Asn Tyr Gln
 210 215 220
 Val Met Ala Thr Glu Gly Tyr Gln Ser Ser Gly Ser Ser Asp Ile Thr
 225 230 235 240
 Val Trp Glu Gly Thr Ser Ser Gly Gly Ser Ser Asn Gly Gly Ser Ser
 245 250 255
 Asn Gly Gly Ser Ser Asn Gly Gly Ser Gly Gly Thr Lys Ser Phe Thr
 260 265 270
 Val Arg Ala Arg Gly Thr Ala Gly Gly Glu Ser Ile Thr Leu Arg Val
 275 280 285
 Asn Asn Gln Asn Val Gln Thr Trp Thr Leu Gly Thr Ser Met Gln Asn
 290 295 300
 Tyr Thr Ala Ser Thr Ser Leu Ser Gly Gly Ile Thr Val Ala Phe Thr
 305 310 315 320
 Asn Asp Gly Gly Ser Arg Asp Val Gln Val Asp Tyr Ile Ile Val Asn
 325 330 335
 Gly Gln Thr Arg Gln Ser Glu Gln Gln Ser Tyr Asn Thr Gly Leu Tyr
 340 345 350
 Ala Asn Gly Ser Cys Gly Gly Gly Ser Asn Ser Glu Trp Met His Cys
 355 360 365
 Asn Gly Ala Ile Gly Tyr Gly Asn Thr Pro
 370 375

<210> 257
 <211> 2694
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample
 Page 185

<400> 257

atggctgata	tatctaccac	accagtcaca	gcctcgacag	atgctgccaa	gaacctgtat	60
gcctattttc	tggaaccagta	tggaagaag	acgattttcca	gcgtcatggc	caatgtcaac	120
tggaacaaca	cttgtgccga	gaaagtctat	aaactcacgg	gcaagtatcc	tgccatgaac	180
tgctacgact	tcatccacat	ctgtttctcg	ccagccaact	ggattgacta	caccgacatc	240
actcctgcca	aggaatggca	cgatgcgggc	ggtatcgtag	agttgatgtg	gcatttcaat	300
gtgcctaaga	gccagggggc	aacagatggt	acctgcacgc	ccagcgagac	cacctttaag	360
gcttccaatg	ctctgggttag	cggcacgtgg	gagaacaaat	ggttctacga	gcagatggac	420
aaggtcattg	ccaccatcct	caagttacag	gacgtggca	ttgccgctac	ctggcgacct	480
ttccatgagg	cagcaggcaa	tgcttgccgc	aagcagcagg	ccgactggac	caaagcatgg	540
ttctggtggg	gctacgacgg	tgccgacacc	tacaagaaac	tgtggattgc	catgtacgac	600
tatttcaagc	tgaaaggcgt	gaacaacctc	atctggatgt	ggaccacca	gaattataat	660
ggtgacagca	gcaaatacaa	ccaggacacc	gactggtagc	ctggcgacga	gtatgttgac	720
atcgtggccc	gcgacctcta	tggtctacaat	gccgaccaga	acctgcagga	gttcagcgag	780
attcaggctg	cctatcccaa	caagatggtg	gttctgggtg	aatgcggaaa	aggtgatagc	840
ggcgaccccg	gcaagatgtc	cgatgtatgg	gcgaaagggtg	ccaagtgggg	ccacttcattg	900
gtatgggtatc	aaggcgaaca	aggctctacc	gacacgatgt	gcagcgacga	ctgggtggaag	960
gatgccatga	gcagcgccaa	cgctatcacc	cgcgacaagg	tggttatccc	cgatgtcact	1020
tcaaccatcg	agaatgccac	ggatgccgtg	aagaacatgg	gactgggggtg	gaacctgggg	1080
aacgcccctcg	acgccaatgc	ccagcaatac	catgatgccca	cccaggacaa	ctactgggga	1140
cagcaggaca	ttacctctga	gagctgctgg	ggtcagctac	ccaccaaggc	agagctgatg	1200
gccatgatga	aagaagccgg	tttcggagcc	atccgcgttc	ccgtgacatg	gtataaccac	1260
atggacaagg	acggcaatgt	ggatgcagca	tggtatgaatc	gtgtgcatga	ggtgggtgac	1320
tatgtcatca	gccagggaat	gtactgcac	ctcaacgtac	accacgacac	gggtgccgac	1380
agctacgaca	gccagaagaa	cctcaccggc	taccattgga	tcaaggccga	cgaaccaaac	1440
tacgccacca	acaaggcccg	ctatgagaag	ctgtggcagc	agatagccca	ggagttccgc	1500
aactacggcc	agctgctgct	gttcgagggc	tataacgaga	tgctcgatgc	caacaactcc	1560
tggaattttg	cacagagcag	ttcagcctac	gatgccatca	acaaatacgc	ccagagcttt	1620
gtcgatgtcg	tacgcgccac	cgggtggcaac	aatgcccagc	gcaacctcat	tgtcagcaca	1680
tacggcgccct	gctcaggcaa	cggcacgtgg	gatgcaagag	tgcaagacc	cttgaagaaa	1740
ctgcagattc	ccacgggtga	aagcaaccat	atcatcttcg	aggttcacia	ctatccctcc	1800
atcgtcaaca	aggacaacgc	gggcaactac	gtcagcgatc	gcaccatcag	cgaaatcaag	1860
gcagagattg	atgcatggct	taagaactta	aagaccaccc	tcgtcagcaa	gggcgctccc	1920
gtcatctatcg	gcgaatgggg	caccaacaac	gtcgatgccg	gcggtggcaa	gacagactac	1980
gacctcata	aggacctgat	gttcgaattt	gtcagctaca	tgataaagac	catgaagcag	2040
aacgacattg	ccaccttcta	ctggatggga	cttaccgacg	gcgctccacg	cacctacccc	2100
gccttcacac	agccccgacct	ggcgctgaag	atgctgcagg	cctatcacgg	cgactcttgg	2160
aatccctacc	tgccctgacgc	caaggacttt	cccgaaggca	aatcacctc	ggccacgggtg	2220
aatttcaaca	gccaatgggg	cgaactgacc	atccacgatg	gagctattga	caagaccgtc	2280
tatagaggta	tcaagggtgga	gctggaagaa	aagcctgcca	ctggagccct	gtctttcaag	2340
gtatatgccca	acagtgaagaa	ggcaacagcc	atcaattcca	aaacccca	gttggctttc	2400
ttcagttaca	caggcatcca	gaaaatcaac	ctacagtgga	acatagccac	caaggggagt	2460
atcaaaaatca	agagcgtcaa	ccttatcaag	cacgacgact	ccacagaacc	ctgtagtctg	2520
aaagtggcctt	ggggttgtag	tctcagcgac	cagaactacg	ccacgggcat	cgaagacatt	2580
actatcactc	ctgttcgtca	tgacgatgga	atcatctaca	atctgagcgg	acagcctgta	2640
acctctctc	agcgcggcat	ctacatcctc	aacggaaaga	aatcatcaa	atag	2694

<210> 258

<211> 897

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample

<400> 258

Met	Ala	Asp	Ile	Ser	Thr	Thr	Pro	Val	Thr	Ala	Ser	Thr	Asp	Ala	Ala
1				5					10				15		
Lys	Asn	Leu	Tyr	Ala	Tyr	Phe	Leu	Asp	Gln	Tyr	Gly	Lys	Lys	Thr	Ile
			20					25				30			
Ser	Ser	Val	Met	Ala	Asn	Val	Asn	Trp	Asn	Asn	Thr	Cys	Ala	Glu	Lys
		35					40				45				
Val	Tyr	Lys	Leu	Thr	Gly	Lys	Tyr	Pro	Ala	Met	Asn	Cys	Tyr	Asp	Phe
	50				55					60					
Ile	His	Ile	Cys	Phe	Ser	Pro	Ala	Asn	Trp	Ile	Asp	Tyr	Thr	Asp	Ile
65				70					75				80		
Thr	Pro	Ala	Lys	Glu	Trp	His	Asp	Ala	Gly	Gly	Ile	Val	Gln	Leu	Met
			85						90				95		

Trp His Phe Asn Val Pro Lys Ser Gln Gly Ala Thr Asp Val Thr Cys
 100 105 110
 Thr Pro Ser Glu Thr Thr Phe Lys Ala Ser Asn Ala Leu Val Ser Gly
 115 120 125
 Thr Trp Glu Asn Lys Trp Phe Tyr Glu Gln Met Asp Lys Val Ile Ala
 130 135 140
 Thr Ile Leu Lys Leu Gln Asp Ala Gly Ile Ala Thr Trp Arg Pro
 145 150 155 160
 Phe His Glu Ala Ala Gly Asn Ala Cys Ala Lys Gln Gln Ala Asp Trp
 165 170 175
 Thr Lys Ala Trp Phe Trp Trp Gly Tyr Asp Gly Ala Asp Thr Tyr Lys
 180 185 190
 Lys Leu Trp Ile Ala Met Tyr Asp Tyr Phe Lys Leu Lys Gly Val Asn
 195 200 205
 Asn Leu Ile Trp Met Trp Thr Thr Gln Asn Tyr Asn Gly Asp Ser Ser
 210 215 220
 Lys Tyr Asn Gln Asp Thr Asp Trp Tyr Pro Gly Asp Glu Tyr Val Asp
 225 230 235 240
 Ile Val Ala Arg Asp Leu Tyr Gly Tyr Asn Ala Asp Gln Asn Leu Gln
 245 250 255
 Glu Phe Ser Glu Ile Gln Ala Ala Tyr Pro Asn Lys Met Val Val Leu
 260 265 270
 Gly Glu Cys Gly Lys Gly Asp Ser Gly Asp Pro Gly Lys Met Ser Asp
 275 280 285
 Val Trp Ala Lys Gly Ala Lys Trp Gly His Phe Met Val Trp Tyr Gln
 290 295 300
 Gly Glu Gln Gly Ser Thr Asp Thr Met Cys Ser Asp Asp Trp Trp Lys
 305 310 315 320
 Asp Ala Met Ser Ser Ala Asn Val Ile Thr Arg Asp Lys Val Val Ile
 325 330 335
 Pro Asp Val Thr Ser Thr Ile Glu Asn Ala Thr Asp Ala Val Lys Asn
 340 345 350
 Met Gly Leu Gly Trp Asn Leu Gly Asn Ala Leu Asp Ala Asn Ala Gln
 355 360 365
 Gln Tyr His Asp Ala Thr Gln Asp Asn Tyr Trp Gly Gln Gln Asp Ile
 370 375 380
 Thr Ser Glu Ser Cys Trp Gly Gln Leu Pro Thr Lys Ala Glu Leu Met
 385 390 395 400
 Ala Met Met Lys Glu Ala Gly Phe Gly Ala Ile Arg Val Pro Val Thr
 405 410 415
 Trp Tyr Asn His Met Asp Lys Asp Gly Asn Val Asp Ala Ala Trp Met
 420 425 430
 Asn Arg Val His Glu Val Val Asp Tyr Val Ile Ser Gln Gly Met Tyr
 435 440 445
 Cys Ile Leu Asn Val His His Asp Thr Gly Ala Asp Ser Tyr Asp Ser
 450 455 460
 Gln Lys Asn Leu Thr Gly Tyr His Trp Ile Lys Ala Asp Glu Thr Asn
 465 470 475 480
 Tyr Ala Thr Asn Lys Ala Arg Tyr Glu Lys Leu Trp Gln Gln Ile Ala
 485 490 495
 Gln Glu Phe Arg Asn Tyr Gly Gln Leu Leu Leu Phe Glu Gly Tyr Asn
 500 505 510
 Glu Met Leu Asp Ala Asn Asn Ser Trp Asn Phe Ala Gln Ser Ser Ser
 515 520 525
 Ala Tyr Asp Ala Ile Asn Lys Tyr Ala Gln Ser Phe Val Asp Val Val
 530 535 540
 Arg Ala Thr Gly Gly Asn Asn Ala Gln Arg Asn Leu Ile Val Ser Thr
 545 550 555 560
 Tyr Gly Ala Cys Ser Gly Asn Gly Thr Trp Asp Ala Arg Val Gln Asp
 565 570 575
 Pro Leu Lys Lys Leu Gln Ile Pro Thr Gly Glu Ser Asn His Ile Ile
 580 585 590
 Phe Glu Val His Asn Tyr Pro Ser Ile Val Asn Lys Asp Asn Ala Gly
 595 600 605
 Asn Tyr Val Ser Asp Arg Thr Ile Ser Glu Ile Lys Ala Glu Ile Asp
 610 615 620
 Ala Trp Leu Lys Asn Leu Lys Thr His Leu Val Ser Lys Gly Ala Pro
 625 630 635 640
 Val Ile Ile Gly Glu Trp Gly Thr Asn Asn Val Asp Ala Gly Gly Gly

Lys Thr Asp Tyr 645
 Tyr Met Ile 660
 Met Gly Leu Thr Asp Gly Ala Pro Arg Thr Tyr 665
 Pro Asp Leu Ala Leu Lys 680
 Asn Pro Tyr Leu Pro Asp Ala Lys Asp Phe Pro Glu Gly Lys 685
 Ser Ala Thr Val 690
 Asp Gly Ala Ile Asp Lys Thr Val Tyr Arg Gly Ile Lys Val Glu Leu 700
 Glu Glu Lys Pro Ala Thr Gly Ala Leu Ser Phe Lys Val Tyr Ala Asn 710
 Ser Glu Lys Ala Thr Ala Ile Asn Ser Lys Thr Pro Glu Leu Ala Phe 715
 Phe Ser Tyr Thr Gly Ile Gln Lys Ile Asn Leu Gln Trp Asn Ile Ala 720
 Thr Lys Gly Ser Ile Lys Ile Lys Ser Val Asn Leu Ile Lys His Asp 725
 Asp Ser Thr Glu Pro Cys Ser Leu Lys Val Ala Trp Gly Cys Thr Leu 730
 Ser Asp Gln Asn Tyr Ala Thr Gly Ile Glu Asp Ile Thr Ile Thr Pro 735
 Val Arg His Asp Asp Gly Ile Ile Tyr Asn Leu Ser Gly Gln Pro Val 740
 Thr Ser Pro Gln Arg Gly Ile Tyr Ile Leu Asn Gly Lys Lys Ile Ile 745
 Lys 750

<210> 259
 <211> 1143
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 259
 atgaagaaaa ttcgcttact ccagggtggt tcggttgcca tgtcaataat gtttcttttg 60
 tcatgtcagg cacaaaaacc agttgactct ctttaagggaag catttgatgg tttgtttctt 120
 ataggtactg ccatgaacac ccctcagatc accggccagg atacacaaac acttgagttg 180
 ataaaaaac acatgaactc catagtggcc gaaaaatgtaa tgaaaagtga ggtgcttcaa 240
 ccagggaag gagagtttga ttttactctt gccgatcagt ttgttcaatt tggatcgtat 300
 aacaatatgc atatagttgg ccataccctt atatggcatt ccaggcgcc acgatggttt 360
 tttgtggatg agaacggaaa cgatgtgagc cccgaaattc tgaaacaaag aatgaaagac 420
 catatttata ccgtagtagg ccgttataaa ggcaaaattc atggatggga tgtggtgaat 480
 gagtgatata atgacgatgg ttcgtggcgc aatagtaagt tttaacaaat tcttggtgaa 540
 gattttgtta aatatgcatt ccagtttgca gctgaagccg atcccgatgc agagctttat 600
 tacaatgatt attcgtatgtt ccttcagga cgtagggaag gcgtaattaa gatgggtgaga 660
 aatctgcagg aacagggaat taaaattgat ggtattggga tgcagggcca cctgatgatt 720
 gattatccac ccctcgaaga ttttgaaacg agtatactgg cttttgccga tctgggggtg 780
 aatgtcatga taaccgaact cgatatatcc gttttgccat ttcctaccgc caacgtgggc 840
 gccgatgttt ctctgaacat tgcatacaat actgaattaa atccctaccg gaatggctta 900
 cccgaagatg tagcgcagaa attacataat cgggtgggtg atctttttcg cctgttcatt 960
 aaacaccacg ataaaattac ccgtgtaacc acttggggta cagccgatgc catgtcatgg 1020
 aagaataact ggcccattcg tggacgtaca gattatccct tacttttcga tcgcatgttt 1080
 cagcccaaac cttttgtcgc tgatataatt aaggaggcat tggcagccaa aagaaaatta 1140
 taa 1143

<210> 260
 <211> 380
 <212> PRT
 <213> Unknown

<220>

<223> Obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(24)

<400> 260

```

Met Lys Lys Ile Arg Leu Leu Gln Gly Val Ser Leu Ala Met Ser Ile
 1      5      10      15
Met Phe Leu Leu Ser Cys Gln Ala Gln Lys Pro Val Asp Ser Leu Lys
      20      25      30
Glu Ala Phe Asp Gly Leu Phe Leu Ile Gly Thr Ala Met Asn Thr Pro
      35      40      45
Gln Ile Thr Gly Gln Asp Thr Gln Thr Leu Glu Leu Ile Lys Lys His
      50      55      60
Met Asn Ser Ile Val Ala Glu Asn Val Met Lys Ser Glu Val Leu Gln
65      70      75      80
Pro Arg Glu Gly Glu Phe Asp Phe Thr Leu Ala Asp Gln Phe Val Gln
      85      90      95
Phe Gly Ile Asp Asn Asn Met His Ile Val Gly His Thr Leu Ile Trp
      100      105      110
His Ser Gln Ala Pro Arg Trp Phe Val Asp Glu Asn Gly Asn Asp
      115      120      125
Val Ser Pro Glu Ile Leu Lys Gln Arg Met Lys Asp His Ile Tyr Thr
      130      135      140
Val Val Gly Arg Tyr Lys Gly Lys Ile His Gly Trp Asp Val Val Asn
145      150      155      160
Glu Cys Ile Asn Asp Asp Gly Ser Trp Arg Asn Ser Lys Phe Tyr Gln
      165      170      175
Ile Leu Gly Glu Asp Phe Val Lys Tyr Tyr Ala Phe Gln Phe Ala Ala Glu
      180      185      190
Ala Asp Pro Asp Ala Glu Leu Tyr Tyr Asn Asp Tyr Ser Met Phe Leu
      195      200      205
Pro Gly Arg Arg Glu Gly Val Ile Lys Met Val Arg Asn Leu Gln Glu
      210      215      220
Gln Gly Ile Lys Ile Asp Gly Ile Gly Met Gln Gly His Leu Met Ile
225      230      235      240
Asp Tyr Pro Pro Leu Glu Asp Phe Glu Thr Ser Ile Leu Ala Phe Ala
      245      250      255
Asp Leu Gly Val Asn Val Met Ile Thr Glu Leu Asp Ile Ser Val Leu
      260      265      270
Pro Phe Pro Thr Arg Asn Val Gly Ala Asp Val Ser Leu Asn Ile Ala
      275      280      285
Tyr Asn Thr Glu Leu Asn Pro Tyr Pro Asn Gly Leu Pro Glu Asp Val
      290      295      300
Ala Gln Lys Leu His Asn Arg Trp Val Asp Leu Phe Arg Leu Phe Ile
305      310      315      320
Lys His His Asp Lys Ile Thr Arg Val Thr Thr Trp Gly Thr Ala Asp
      325      330      335
Ala Met Ser Trp Lys Asn Asn Trp Pro Ile Arg Gly Arg Thr Asp Tyr
      340      345      350
Pro Leu Leu Phe Asp Arg Asp Phe Gln Pro Lys Pro Phe Val Ala Asp
      355      360      365
Ile Ile Lys Glu Ala Leu Ala Ala Lys Arg Lys Leu
      370      375      380

```

<210> 261

<211> 1629

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 261

```

atgataaaca aaattggcaa aggttttttt tctgcttca tttgtgctgc tgcgttgagt
gtctccacag ttaatgctca gcaaactgtc accaccaaca cgcaaggcac gcacgatggt
ttttctatt cgttttggaa agacagtggg gatgcatcat ttggtttgcg tgagggaggg
cgttacacct cgcaatggaa tacttctacc aataactggg tgggtggaaa aggggtggaat

```

```

60
120
180
240

```


cccgggtgga	gaaggggtgt	tcactatcaa	ggccaatata	atgttgataa	ttcacaaaac	300
tcttatttgg	cattgtatgg	ctggacacgc	tcaccactga	ttgaatatta	cgtgattgaa	360
agttacggct	cgtataaccc	gtcgaattgc	acccaaggct	ggcagaccta	tggcaccctt	420
cagagtgatg	gtgcaaccta	tgaattgttt	cgctgtcagc	gagttcagca	gccctctatc	480
gatggcacac	aaacttttcta	tcaatacttc	agtgtgcgtc	agccgaagaa	aggctttggt	540
agtatcagtg	gtacgatcac	tgtgggcaac	cattttgatg	catggggccg	cggcgggttg	600
aacctggggg	aacatgatta	tatgggtgat	gctaccgagg	gttatcagag	caccggtagt	660
tcggatatta	cggtcagtga	aattaccggg	ggttcagggt	gtggctcttc	ctcgggtgct	720
aataccctgg	tgattcgtgc	tgtgggcacc	tctggtaatg	aattgctgcg	tgtcaatgtg	780
ggtaggtgac	ctgtgcagac	attgagcctt	tcgaccagtt	ggcaggattt	tactgtcaat	840
acggatgcaa	cgggtgacat	taacgtagag	ttgtttaatg	atcaggggtca	gggttatgag	900
gcgcgtatcg	attatgtgct	ggttaatggg	gagaccgcgt	acgcggccga	tcagagttat	960
aacaccagtg	cctgggacgg	cgaatgtggg	ggtagcctct	ttaccagtg	gatgcattgt	1020
gatggcatga	ttggctttgg	tgatatgacc	ggcggcaatg	ccggtgggtg	cggttcttcg	1080
ggtaggtctg	gcgcgaatac	tctggtgggt	cgctgtgtcg	gcacttcagg	taacgagcag	1140
ttgcgcgtga	atgtgggcgg	caacacgatt	caaactctga	acctgtcaag	cagttggcaa	1200
gattttactg	tcaataccga	tgccctcggg	gatattaacg	tagagctgtt	taatgaccag	1260
ggtaggggct	atgaggcgcg	tattgattat	gtgctgggta	atggcgagac	ccgctacgcg	1320
gctgaccaga	gttataaacac	cagcgcctgg	gatggcgaat	gcgggggtgg	ctcttttacc	1380
caatggatgc	attgtgatgg	catgattggg	tttgggtgata	tgtcgggtgg	tggttctgct	1440
gtgggtacaa	tgatgacggg	taatgccggc	agcaataacca	gcagtgcctg	ttactgtaat	1500
tggtatggca	gtgtgatggc	ttcttgtgaa	aatcaggtga	acggctgggg	ttgggaaaat	1560
aatcaaagct	gtattggtaa	taatacctgt	aataatcagg	gcggtagcgg	aggcgtgggt	1620
tgcaattaa						1629

<210> 262

<211> 542

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(26)

<400> 262

Met	Ile	Asn	Lys	Ile	Gly	Lys	Gly	Phe	Phe	Ser	Ala	Phe	Ile	Cys	Ala
1				5				10						15	
Ala	Ala	Leu	Ser	Val	Ser	Thr	Val	Asn	Ala	Gln	Gln	Thr	Val	Thr	Thr
			20					25					30		
Asn	Thr	Gln	Gly	Thr	His	Asp	Gly	Phe	Phe	Tyr	Ser	Phe	Trp	Lys	Asp
		35					40					45			
Ser	Gly	Asp	Ala	Ser	Phe	Gly	Leu	Arg	Glu	Gly	Gly	Arg	Tyr	Thr	Ser
	50					55					60				
Gln	Trp	Asn	Thr	Ser	Thr	Asn	Asn	Trp	Val	Gly	Gly	Lys	Gly	Trp	Asn
65					70			75						80	
Pro	Gly	Gly	Arg	Arg	Val	Val	His	Tyr	Gln	Gly	Gln	Tyr	Asn	Val	Asp
			85					90					95		
Asn	Ser	Gln	Asn	Ser	Tyr	Leu	Ala	Leu	Tyr	Gly	Trp	Thr	Arg	Ser	Pro
			100					105					110		
Leu	Ile	Glu	Tyr	Tyr	Val	Ile	Glu	Ser	Tyr	Gly	Ser	Tyr	Asn	Pro	Ser
		115					120					125			
Asn	Cys	Thr	Gln	Gly	Arg	Gln	Thr	Tyr	Gly	Thr	Phe	Gln	Ser	Asp	Gly
	130					135					140				
Ala	Thr	Tyr	Glu	Ile	Val	Arg	Cys	Gln	Arg	Val	Gln	Gln	Pro	Ser	Ile
145					150					155					160
Asp	Gly	Thr	Gln	Thr	Phe	Tyr	Gln	Tyr	Phe	Ser	Val	Arg	Gln	Pro	Lys
			165						170					175	
Lys	Gly	Phe	Gly	Ser	Ile	Ser	Gly	Thr	Ile	Thr	Val	Gly	Asn	His	Phe
		180						185					190		
Asp	Ala	Trp	Ala	Ala	Ala	Gly	Leu	Asn	Leu	Gly	Glu	His	Asp	Tyr	Met
	195						200					205			
Val	Met	Ala	Thr	Glu	Gly	Tyr	Gln	Ser	Thr	Gly	Ser	Ser	Asp	Ile	Thr
	210					215					220				
Val	Ser	Glu	Ile	Thr	Gly	Gly	Ser	Gly	Gly	Gly	Ser	Ser	Ser	Gly	Ala
225					230					235					240
Asn	Thr	Leu	Val	Ile	Arg	Ala	Val	Gly	Thr	Ser	Gly	Asn	Glu	Leu	Leu
			245					250					255		

Arg Val Asn Val Gly Gly Ser Pro Val Gln Thr Leu Ser Leu Ser Thr
 260 265 270
 Ser Trp Gln Asp Phe Thr Val Asn Thr Asp Ala Thr Gly Asp Ile Asn
 275 280 285
 Val Glu Leu Phe Asn Asp Gln Gly Gln Gly Tyr Glu Ala Arg Ile Asp
 290 295 300
 Tyr Val Leu Val Asn Gly Glu Thr Arg Tyr Ala Asp Gln Ser Tyr
 305 310 315 320
 Asn Thr Ser Ala Trp Asp Gly Glu Cys Gly Gly Gly Ser Phe Thr Gln
 325 330 335
 Trp Met His Cys Asp Gly Met Ile Gly Phe Gly Asp Met Thr Gly Gly
 340 345 350
 Asn Ala Gly Gly Gly Gly Ser Ser Gly Gly Ser Gly Ala Asn Thr Leu
 355 360 365
 Val Val Arg Ala Val Gly Thr Ser Gly Asn Glu Gln Leu Arg Val Asn
 370 375 380
 Val Gly Gly Asn Thr Ile Gln Thr Leu Asn Leu Ser Ser Ser Trp Gln
 385 390 395 400
 Asp Phe Thr Val Asn Thr Asp Ala Ser Gly Asp Ile Asn Val Glu Leu
 405 410 415
 Phe Asn Asp Gln Gly Gln Gly Tyr Glu Ala Arg Ile Asp Tyr Val Leu
 420 425 430
 Val Asn Gly Glu Thr Arg Tyr Ala Ala Asp Gln Ser Tyr Asn Thr Ser
 435 440 445
 Ala Trp Asp Gly Glu Cys Gly Gly Ser Phe Thr Gln Trp Met His
 450 455 460
 Cys Asp Gly Met Ile Gly Phe Gly Asp Met Ser Gly Gly Gly Ser Ala
 465 470 475 480
 Val Gly Thr Ser Ser Gly Asn Ala Gly Ser Asn Thr Ser Ser Ala
 485 490 495
 Cys Tyr Cys Asn Trp Tyr Gly Ser Val Met Ala Ser Cys Glu Asn Gln
 500 505 510
 Val Asn Gly Trp Gly Trp Glu Asn Asn Gln Ser Cys Ile Gly Asn Asn
 515 520 525
 Thr Cys Asn Asn Gln Gly Gly Ser Gly Gly Val Val Cys Asn
 530 535 540

<210> 263
 <211> 1092
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 263
 atgaaaacta atcacccatt taaattcggg aaaaaaatat gtatggcatt ggctttgctg 60
 gtgcttggca tacaggcttc aatcgcacag gaaatttgta ttaccagcgg cactgaccag 120
 atcagagaaa ccacatccaa cggctataacc cactgaactat ggaatcagga caccggggg 180
 acggcctgta tgactattaa tgcaggcacc acttacagtg cgcggtggaa cgggtgcattt 240
 aactatttgg cccgccgtgg attggcctac gatgggttcgt ccctcaccca tgctgaccgg 300
 gggaaaattca ccataaatta tgcctctaac tacaactgca acaatatgaa tgggctctct 360
 tatttaagcg tgtacggatg gacgcgggat ttggccaagg aaaatgcaa tccggcagga 420
 tcacaggctc atcaggaagc gctgggtggaa tattacattg ttgaaaactg gtgcgactgg 480
 aatgtttcac aagaccctaa cgcccagagt ctgggcaccc tgaatgttga tgggtcgatc 540
 tatgatatgt atcgcacaga acggatcaac caaccttcta tcagggtgcgg tgggtacctgc 600
 gataattttt accaatactt cagcattcgc cgcaacacac gtaacagtgg caccattgat 660
 gtcagcgctc atttcaacca gtgggaagca ttaaccggcg tccctatggg tggcctgcac 720
 gaagtgatga tgaaggtcga aggtacaac tcaaacaatc aatccagtgg caatgtaagc 780
 tttactcaat tgctcatgcg tgcccgttc gaggatggcg ccattgtcga gaaccagaat 840
 gcggtcggcc atgcgcacgg tggagaagcg gtgggagatg atcaccgccg tcttgccctg 900
 gcccaggccc ttgaagcggg cgaacacctc ggccctcgcc ttggcgtcga gggcggcggt 960
 gggttcgtcg agaattgatca actcggcgctc gcgcgtatag tggctacctt 1020
 ctgccactcg ccgcccagaga ggtcgcggcc ctgcttgaaa aggcgccccca attgctccag 1080
 agaaacgggt ga 1092

<210> 264
 <211> 363
 <212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(29)

<400> 264

```

Met Lys Thr Asn His Pro Phe Lys Phe Gly Lys Lys Ile Cys Met Ala
 1      5      10      15
Leu Ala Leu Leu Val Leu Gly Ile Gln Ala Ser Ile Ala Gln Glu Ile
 20      25      30
Cys Ile Thr Ser Gly Thr Asp Gln Ile Arg Glu Thr Thr Ser Asn Gly
 35      40      45
Tyr Thr His Glu Leu Trp Asn Gln Asp Thr Arg Gly Thr Ala Cys Met
 50      55      60
Thr Ile Asn Ala Gly Thr Thr Tyr Ser Ala Arg Trp Asn Gly Ala Phe
 65      70      75      80
Asn Tyr Leu Ala Arg Arg Gly Leu Ala Tyr Asp Gly Ser Ser Leu Thr
 85      90      95
His Ala Asp Arg Gly Lys Phe Thr Ile Asn Tyr Ala Ser Asn Tyr Asn
100      105      110
Cys Asn Asn Met Asn Gly Leu Ser Tyr Leu Ser Val Tyr Gly Trp Thr
115      120      125
Arg Asp Phe Ala Lys Glu Asn Ala Asn Pro Ala Gly Ser Gln Ala His
130      135      140
Gln Glu Ala Leu Val Glu Tyr Tyr Ile Val Glu Asn Trp Cys Asp Trp
145      150      155      160
Asn Val Ser Gln Asp Pro Asn Ala Gln Ser Leu Gly Thr Leu Asn Val
165      170      175
Asp Gly Ser Ile Tyr Asp Met Tyr Arg Thr Glu Arg Ile Asn Gln Pro
180      185      190
Ser Ile Arg Cys Gly Gly Thr Cys Asp Asn Phe Tyr Gln Tyr Phe Ser
195      200      205
Ile Arg Arg Asn Thr Arg Asn Ser Gly Thr Ile Asp Val Ser Ala His
210      215      220
Phe Asn Gln Trp Glu Ala Leu Thr Gly Val Pro Met Gly Gly Leu His
225      230      235      240
Glu Val Met Met Lys Val Glu Gly Tyr Asn Ser Asn Asn Gln Ser Ser
245      250      255
Gly Asn Val Ser Phe Thr Gln Leu Leu Met Arg Ala Arg Phe Glu Asp
260      265      270
Gly Ala Ile Val Glu Asn Gln Asn Ala Val Gly His Ala His Gly Gly
275      280      285
Glu Ala Val Gly Asp Asp His Arg Arg Leu Ala Leu Gly Gln Ala Leu
290      295      300
Glu Ala Gly Glu His Leu Gly Leu Gly Leu Val Glu Gly Gly Gly
305      310      315      320
Gly Phe Val Glu Asn Asp Gln Leu Gly Val Ala His Ile Gly Ala Gly
325      330      335
Asp Gly Tyr Leu Leu Pro Leu Ala Ala Arg Glu Val Ala Ala Leu Leu
340      345      350
Glu Lys Ala Pro Gln Leu Leu Gln Arg Asn Gly
355      360

```

<210> 265

<211> 996

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 265

```

atgaacagct ccctccctc cctccgcgat gtattcgcga atgatttcgc catcggggcg      60
gcggtcaatc ctgtgacgat cgagatgcaa aaacagttgt tgatcgatca tgtcaacagt      120
attacggcag agaaccatat gaagtttgag catcttcagc cggaagaagg gaaatttacc      180
tttcaggaag cggatcggat tgtggatitt gcttggttcgc accgaatggc ggttcgaggg      240

```

cacacacttg	tatggcacia	ccagactccg	gattgggtgt	ttcaagatgg	tcaaggccat	300
ttcgtcagtc	gggatgtgtt	gcttgagcgg	atgaaatgtc	acatttcaac	tggtgtacgg	360
cgatacaagg	gaaaaatata	ttgttgggat	gtcatcaacg	aagcggtagc	cgacgaagga	420
gacgaattgt	tgaggccgctc	gaagtggcga	caaatcatcg	gggacgattt	tatggaacaa	480
gcatttctct	acgcttatga	agctgaccca	gatgcactgc	ttttttacaa	tgactataat	540
gaatgttttc	cggaaaagag	agaaaaaatt	tttgacttg	tcaaatcgct	gcgtgataaa	600
ggcattccga	ttcatggcat	cggcatgcag	gcgcactgga	gcctgaccg	cccgtcgctt	660
gatgaaattc	gtgcggcgat	tgaacgggat	gcgtcccttg	gtgttgttct	tcatattacg	720
gaactcgatg	tatccatggt	tgaatttcac	gatcgtcgaa	ccgatttggc	tgtcccgcacg	780
aacgaaatga	tcgaacagca	agcagaacgg	tatgggcaaa	tttttgcttt	gtttaaggag	840
tatcgcgatg	ttattcaaag	tgtcacattt	tggggaattg	ctgatgacca	tacatggctc	900
gataactttc	cagtgcacgg	gagaaaaaac	tggccgcttt	tgttcgatga	acagcataaa	960
ccgaaaccag	ctttttggcg	ggcagtgagt	gtctga			996

<210> 266

<211> 331

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 266

Met	Asn	Ser	Ser	Leu	Pro	Ser	Leu	Arg	Asp	Val	Phe	Ala	Asn	Asp	Phe
1				5					10				15		
Arg	Ile	Gly	Ala	Ala	Val	Asn	Pro	Val	Thr	Ile	Glu	Met	Gln	Lys	Gln
			20					25					30		
Leu	Leu	Ile	Asp	His	Val	Asn	Ser	Ile	Thr	Ala	Glu	Asn	His	Met	Lys
			35					40				45			
Phe	Glu	His	Leu	Gln	Pro	Glu	Gly	Lys	Phe	Thr	Phe	Gln	Glu	Ala	
	50					55				60					
Asp	Arg	Ile	Val	Asp	Phe	Ala	Cys	Ser	His	Arg	Met	Ala	Val	Arg	Gly
65					70				75					80	
His	Thr	Leu	Val	Trp	His	Asn	Gln	Thr	Pro	Asp	Trp	Val	Phe	Gln	Asp
			85						90				95		
Gly	Gln	Gly	His	Phe	Val	Ser	Arg	Asp	Val	Leu	Leu	Glu	Arg	Met	Lys
			100					105					110		
Cys	His	Ile	Ser	Thr	Val	Val	Arg	Tyr	Lys	Gly	Lys	Ile	Tyr	Cys	
			115				120				125				
Trp	Asp	Val	Ile	Asn	Glu	Ala	Val	Ala	Asp	Glu	Gly	Asp	Glu	Leu	Leu
	130					135				140					
Arg	Pro	Ser	Lys	Trp	Arg	Gln	Ile	Ile	Gly	Asp	Asp	Phe	Met	Glu	Gln
145					150				155					160	
Ala	Phe	Leu	Tyr	Ala	Tyr	Glu	Ala	Asp	Pro	Asp	Ala	Leu	Leu	Phe	Tyr
			165						170					175	
Asn	Asp	Tyr	Asn	Glu	Cys	Phe	Pro	Glu	Lys	Arg	Glu	Lys	Ile	Phe	Ala
			180					185					190		
Leu	Val	Lys	Ser	Leu	Arg	Asp	Lys	Gly	Ile	Pro	Ile	His	Gly	Ile	Gly
		195					200					205			
Met	Gln	Ala	His	Trp	Ser	Leu	Thr	Arg	Pro	Ser	Leu	Asp	Glu	Ile	Arg
	210					215				220					
Ala	Ala	Ile	Glu	Arg	Tyr	Ala	Ser	Leu	Gly	Val	Val	Leu	His	Ile	Thr
225					230				235						240
Glu	Leu	Asp	Val	Ser	Met	Phe	Glu	Phe	His	Asp	Arg	Arg	Thr	Asp	Leu
			245						250					255	
Ala	Val	Pro	Thr	Asn	Glu	Met	Ile	Glu	Gln	Gln	Ala	Glu	Arg	Tyr	Gly
			260					265					270		
Gln	Ile	Phe	Ala	Leu	Phe	Lys	Glu	Tyr	Arg	Asp	Val	Ile	Gln	Ser	Val
		275					280					285			
Thr	Phe	Trp	Gly	Ile	Ala	Asp	Asp	His	Thr	Trp	Leu	Asp	Asn	Phe	Pro
	290					295					300				
Val	His	Gly	Arg	Lys	Asn	Trp	Pro	Leu	Leu	Phe	Asp	Glu	Gln	His	Lys
305					310					315					320
Pro	Lys	Pro	Ala	Phe	Trp	Arg	Ala	Val	Ser	Val					
			325						330						

<210> 267

<211> 1956

<212> DNA

<213> Bacteria

<400> 267

atgaagcgta	aggттааgaa	gatggcagct	atggcaacga	gtataattat	ggctatcatg	60
atcatcctac	atagtatacc	agtactcgcc	gggcgaataa	tttacgacaa	tgagacaggc	120
acacatggag	gctacgacta	tgagctctgg	aaagactacg	gaaatacgat	tatggaactt	180
aacgacgggtg	gtacttttag	ttgtcaatgg	agtaatatcg	gtaatgcact	atttagaaaa	240
gggagaaaaat	ttaattccga	caaaacctat	caagaattag	gagatatagt	agttgaatat	300
ggctgtgatt	acaatccaaa	cggaaattcc	tatttgtgtg	tttacggttg	gacaagaaat	360
ccactggttg	aatattacat	tgtagaaagc	tggggcagct	ggcgtccacc	tgagcaaca	420
cccaaaggaa	ccatcacagt	ggatggcggg	acttatgaaa	tatatgaaac	taccgggta	480
aatcagcctt	ccatcgatgg	aactgcgaca	ttccaacaat	attggagtgt	tcgtacatcc	540
aagagaacaa	cgggaacaat	atctgtcact	gaacatttta	aacagtggga	aagaatgggc	600
atgcgaatgg	gtaagatgta	tgaagttgct	cttaccgttg	aaggttatca	gagcagtggtg	660
tacgctaattg	tatataagaa	tgaatcaga	ataggtgcaa	atccaactcc	tgccccatct	720
caaagcccaa	ttagaagaga	tgcattttca	ataatcgaag	cggagaata	taacagcaca	780
aattcctcca	ctttacaagt	gatttgaacg	ccaaataatg	gcagaggaat	tggttatatt	840
gaaaatggta	ataccgtaac	ttacagcaat	atagattttg	gtagtgggtg	aacagggttc	900
tctgcaactg	ttgcaacgga	ggttaatacc	tcaattcaaa	tccgttctga	cagtcctatc	960
ggaactctac	ttggtacctt	atatgtaagt	tctaccggca	gctggaatac	atatcaaacc	1020
gtatctacaa	acatcagcaa	aattaccggc	gttcatgata	ttgtattggt	attctcaggt	1080
ccagtcaatg	tggaacaact	catattttagc	agaagttcac	cagtgcctgc	acctggtgat	1140
aacacaagag	acgcatattc	tatcattcag	gccgaggatt	atgacagcag	ttatggcccc	1200
aaccttcaaa	tccttagctt	accaggcggt	ggcagcgcca	ttggctatat	tgaaaatggt	1260
tattccacta	cctataataa	cgttaatctc	gccaacggct	taagttctat	aacagcaaga	1320
gttgccactc	agatctcaac	ttccattcag	gtgagagcag	gaggagcaac	cggtacttta	1380
cttggtacaa	tatatgttcc	ttcgacaaat	agttgggatt	cttatcagaa	tgtaactgcc	1440
aaccttagca	atattacagg	tgtgcatgat	attacccttg	tcttttcagg	accagtgaat	1500
gtggactact	tcgtattttac	accagcaaat	gtaaatccag	ggcctaccctc	ccctgtcgga	1560
ggtacaagaa	gtgcattttc	caatattcaa	gccgaagatt	atgacagcag	ttatgggtccc	1620
aaccttcaaa	tccttagctt	accagggtggt	ggcagcgcca	ttggctatat	tgaaaatggt	1680
tattccacta	cctataaaaa	tattgatttt	ggtgacggcg	caacgtccgt	aacagcaaga	1740
gtagctaccc	agaatgctac	taccattcag	gtaagattgg	gaagtccatc	gggtacatta	1800
cttgggaacaa	tttacgtggg	gtccacagga	agctttgata	cttataggga	tgtatccgct	1860
accattagta	atactgcggg	tgtaaaagat	attgttcttg	tattttcagg	tcctgttaat	1920
gttgactggg	ttgtattctc	aaaatcagga	acttaa			1956

<210> 268

<211> 651

<212> PRT

<213> Bacteria

<220>

<221> SIGNAL

<222> (1)...(30)

<400> 268

Met	Lys	Arg	Lys	Val	Lys	Lys	Met	Ala	Ala	Met	Ala	Thr	Ser	Ile	Ile
1				5				10					15		
Met	Ala	Ile	Met	Ile	Ile	Leu	His	Ser	Ile	Pro	Val	Leu	Ala	Gly	Arg
			20					25					30		
Ile	Ile	Tyr	Asp	Asn	Glu	Thr	Gly	Thr	His	Gly	Gly	Tyr	Asp	Tyr	Glu
		35				40						45			
Leu	Trp	Lys	Asp	Tyr	Gly	Asn	Thr	Ile	Met	Glu	Leu	Asn	Asp	Gly	Gly
	50					55					60				
Thr	Phe	Ser	Cys	Gln	Trp	Ser	Asn	Ile	Gly	Asn	Ala	Leu	Phe	Arg	Lys
65				70					75				80		
Gly	Arg	Lys	Phe	Asn	Ser	Asp	Lys	Thr	Tyr	Gln	Glu	Leu	Gly	Asp	Ile
			85					90					95		
Val	Val	Glu	Tyr	Gly	Cys	Asp	Tyr	Asn	Pro	Asn	Gly	Asn	Ser	Tyr	Leu
			100					105					110		
Cys	Val	Tyr	Gly	Trp	Thr	Arg	Asn	Pro	Leu	Val	Glu	Tyr	Tyr	Ile	Val
			115				120					125			
Glu	Ser	Trp	Gly	Ser	Trp	Arg	Pro	Pro	Gly	Ala	Thr	Pro	Lys	Gly	Thr
	130					135					140				
Ile	Thr	Val	Asp	Gly	Gly	Thr	Tyr	Glu	Ile	Tyr	Glu	Thr	Thr	Arg	Val
145					150					155				160	
Asn	Gln	Pro	Ser	Ile	Asp	Gly	Thr	Ala	Thr	Phe	Gln	Gln	Tyr	Trp	Ser
				165				170						175	

Val Arg Thr Ser Lys Arg Thr Ser Gly Thr Ile Ser Val Thr Glu His
 Phe Lys Gln Trp Glu Arg Met Gly Met Arg Met Gly Lys Met Tyr Glu
 Val Ala Leu Thr Val Glu Gly Tyr Gln Ser Ser Gly Tyr Ala Asn Val
 Tyr Lys Asn Glu Ile Arg Ile Gly Ala Asn Pro Thr Pro Ala Pro Ser
 Gln Ser Pro Ile Arg Arg Asp Ala Phe Ser Ile Ile Glu Ala Glu Glu
 Tyr Asn Ser Thr Asn Ser Ser Thr Leu Gln Val Ile Gly Thr Pro Asn
 Asn Gly Arg Gly Ile Gly Tyr Ile Glu Asn Gly Asn Thr Val Thr Tyr
 Ser Asn Ile Asp Phe Gly Ser Gly Ala Thr Gly Phe Ser Ala Thr Val
 Ala Thr Glu Val Asn Thr Ser Ile Gln Ile Arg Ser Asp Ser Pro Ile
 Gly Thr Leu Leu Gly Thr Leu Tyr Val Ser Ser Thr Gly Ser Trp Asn
 Thr Tyr Gln Thr Val Ser Thr Asn Ile Ser Lys Ile Thr Gly Val His
 Asp Ile Val Leu Val Phe Ser Gly Pro Val Asn Val Asp Asn Phe Ile
 Phe Ser Arg Ser Ser Pro Val Pro Ala Pro Gly Asp Asn Thr Arg Asp
 Ala Tyr Ser Ile Ile Gln Ala Glu Asp Tyr Asp Ser Ser Tyr Gly Pro
 Asn Leu Gln Ile Phe Ser Leu Pro Gly Gly Gly Ser Ala Ile Gly Tyr
 Ile Glu Asn Gly Tyr Ser Thr Thr Tyr Asn Asn Val Asn Phe Ala Asn
 Gly Leu Ser Ser Ile Thr Ala Arg Val Ala Thr Gln Ile Ser Thr Ser
 Ile Gln Val Arg Ala Gly Gly Ala Thr Gly Thr Leu Leu Gly Thr Ile
 Tyr Val Pro Ser Thr Asn Ser Trp Asp Ser Tyr Gln Asn Val Thr Ala
 Asn Leu Ser Asn Ile Thr Gly Val His Asp Ile Thr Leu Val Phe Ser
 Gly Pro Val Asn Val Asp Tyr Phe Val Phe Thr Pro Ala Asn Val Asn
 Ser Gly Pro Thr Ser Pro Val Gly Gly Thr Arg Ser Ala Phe Ser Asn
 Ile Gln Ala Glu Asp Tyr Asp Ser Ser Tyr Gly Pro Asn Leu Gln Ile
 Phe Ser Leu Pro Gly Gly Gly Ser Ala Ile Gly Tyr Ile Glu Asn Gly
 Tyr Ser Thr Thr Tyr Lys Asn Ile Asp Phe Gly Asp Gly Ala Thr Ser
 Val Thr Ala Arg Val Ala Thr Gln Asn Ala Thr Thr Ile Gln Val Arg
 Leu Gly Ser Pro Ser Gly Thr Leu Leu Gly Thr Ile Tyr Val Gly Ser
 Thr Gly Ser Phe Asp Thr Tyr Arg Asp Val Ser Ala Thr Ile Ser Asn
 Thr Ala Gly Val Lys Asp Ile Val Leu Val Phe Ser Gly Pro Val Asn
 Val Asp Trp Phe Val Phe Ser Lys Ser Gly Thr
 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650

<210> 269
 <211> 1110
 <212> DNA
 <213> unknown

<220>
 <223> obtained from an environmental sample.

<400> 269
 atgggggtaca ataggatcat acaagcgatc cgcgtaagca agggagatgt tttgggcgtt 60
 cataaagttt tttacgtgc acttgctgt gtggcgatgg ggtattcgga aacgtgggca 120
 cagtgcgcga cctggacccg aagcaccatt cgcaattgcg agggcatcga ctacgagttg 180
 tggaaccaga acaaccgcgg cacggtcaac atggaaatca cgggaaacgg aacgttcgcg 240
 gcgacgtgga gcggaacgga aaacatcctg tttcgcgcgg gcaagaaatg ggggttcaac 300
 agcaccacga cggcgcggtc ggtcggcgcc atcacgctcg atttcgctgc gacctggacc 360
 tccagcgaca acgtgaaaat gctcggcatc tacggctggg cgtattaccg gtcgggaagc 420
 gagccgacga aaacggaaag cggtaaaaac acgagctttt ccgatcagat cgagtattac 480
 atcatccagg accgcggagg cttcaacccg ggttcggcg gcgtcaacgc caaaaagtac 540
 ggcgaggccg tgatcgacgg aatcgccat gacttttggg tggccgaccg gatcaaccag 600
 cccatgctga caggaagagg caacttcaag caatacttca gcgttcacg gaacacgagc 660
 agccaccggc aaagcggcat cgtcagcatt tcgaagcact ttgaggagt ggacaaggcc 720
 ggcatgaaga tgctggactg tccgtatac gaagtgcga tgaaggtgga atcgatac 780
 ggctcggcga atggcggcgg gtcggcgaac gtgacccgga atattctcac gctcggcgg 840
 tcttcgcac gcaccctat cgcgcgcggc cccggccggg ccgccgaaag catgcgggtc 900
 gccttcgttc aggaagagt gctcaagggt gcgcccgtcg acggaaccg cctgcaagtc 960
 aaggtgcggg acgtgaagg cgtgaaccgg gccgagttca atgccgcgg cgcggaacg 1020
 ttctcgttgt cccatgtccc cgcgggcccg tatttcctgg atgtgacgg gccggatgta 1080
 agacagatca cgccgttcgt tttgcgataa 1110

<210> 270

<211> 369

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 270

Met Gly Tyr Asn Arg Ile Ile Gln Ala Ile Arg Val Ser Lys Gly Asp
 1 5 10 15
 Val Leu Gly Val His Lys Val Phe Tyr Ala Ala Leu Ala Cys Val Ala
 20 25 30
 Met Gly Tyr Ser Glu Thr Trp Ala Gln Cys Ala Thr Trp Thr Arg Ser
 35 40 45
 Thr Ile Arg Asn Cys Glu Gly Ile Asp Tyr Glu Leu Trp Asn Gln Asn
 50 55 60
 Asn Arg Gly Thr Val Asn Met Glu Ile Thr Gly Asn Gly Thr Phe Ala
 65 70 75 80
 Ala Thr Trp Ser Gly Thr Glu Asn Ile Leu Phe Arg Ala Gly Lys Lys
 85 90 95
 Trp Gly Phe Asn Ser Thr Thr Thr Ala Arg Ser Val Gly Ala Ile Thr
 100 105 110
 Leu Asp Phe Ala Ala Thr Trp Thr Ser Ser Asp Asn Val Lys Met Leu
 115 120 125
 Gly Ile Tyr Gly Trp Ala Tyr Tyr Pro Ser Gly Ser Glu Pro Thr Lys
 130 135 140
 Thr Glu Ser Gly Gln Asn Thr Ser Phe Ser Asp Gln Ile Glu Tyr Tyr
 145 150 155 160
 Ile Ile Gln Asp Arg Gly Gly Phe Asn Pro Gly Ser Gly Gly Val Asn
 165 170 175
 Ala Lys Lys Tyr Gly Glu Ala Val Ile Asp Gly Ile Ala Tyr Asp Phe
 180 185 190
 Trp Val Ala Asp Arg Ile Asn Gln Pro Met Leu Thr Gly Arg Gly Asn
 195 200 205
 Phe Lys Gln Tyr Phe Ser Val Pro Arg Asn Thr Ser Ser His Arg Gln
 210 215 220
 Ser Gly Ile Val Ser Ile Ser Lys His Phe Glu Glu Trp Asp Lys Ala
 225 230 235 240
 Gly Met Lys Met Leu Asp Cys Pro Leu Tyr Glu Val Ala Met Lys Val
 245 250 255
 Glu Ser Tyr Thr Gly Ser Ala Asn Gly Gly Gly Ser Ala Asn Val Thr
 260 265 270
 Arg Asn Ile Leu Thr Leu Gly Gly Ser Ser Ala Pro Thr Pro Ile Ala
 275 280 285
 Arg Gly Pro Gly Arg Ser Ala Glu Ser Met Arg Val Ala Phe Val Gln
 290 295 300
 Glu Arg Val Leu Lys Val Ala Pro Val Asp Gly Thr Arg Leu Gln Val

305 310 315 320
 Lys Val Arg Asp Val Lys Gly Val Asn Arg Ala Glu Phe Asn Ala Ala
 325 330 335
 Gly Ala Ala Thr Phe Ser Leu Ser His Val Pro Ala Gly Pro Tyr Phe
 340 345 350
 Leu Asp Val Thr Gly Pro Asp Val Arg Gln Ile Thr Pro Phe Val Leu
 355 360 365
 Arg

<210> 271
 <211> 1128
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 271
 atgttcattc acaacagcat atgcagcgca ctctgcacaa tctttttggc aactgcaaca 60
 atgggagaaa acatgacact acaagaagcc ttgtccgatc acttttatgt gggagccgcc 120
 atcagccaac gcctttttca accagatcgc gccgaaacgc tgcaactggc cgcgcaccaa 180
 ttcaacagca tcacagccga aaatgagatg aagtggcagt cgtaaatacc cactcctggc 240
 gaataccgtt tcgaaaacgc cgataaattc gtccgccttg gtgtcgaaaa cgatatgtac 300
 atcgttgggc acgttctctt ctggcacagc cagacacccg actggctctt caaggatgac 360
 gacggttaact tcgtctcccg cgaagtctta ctcgaccgca tgcgcgcca cggtgcgaat 420
 cttgtccagc gctacggcaa ccatgtgcac gcctgggatg ttatcaatga aaccttcaat 480
 gataatgggt ccttgcgcga cagcccatgg acgcaaatcc tcggcgagga attcatcgag 540
 cacgccttcc ggattgccgg cgaggaactc ccccccatg tcgagctgct ctacaatgat 600
 tattcgatga ccattcctgc caagcgcgat gctgttgcgt aaatgggtcg cgacctcata 660
 gccaaaggca tccgcattga cggcgttggc atgcagggac attgggcacg gaccacccg 720
 accatagcgg acatagaaaa aagcattctt gccttcgcag gaaccggcgt acaggtacac 780
 atcactgagc tcgacatcga catgctgccca cgccatcccc agatgtttac tgggtggtgca 840
 gacaccatgt tgcgcctaca acaagatccc aaactcgacc cctacactga gggacttcca 900
 gcggaagatc agcaggcatt ggcagaacgc tacgcaagca tcttcggtt attcttgaag 960
 cacagcgatg ttattcgccg tgtcaccttc tgggggggtca ccgatgccca cacctggctc 1020
 aacaattggc ccatccgtgg ccgcaccagc catcccctgc tcttcgaccg ccagaacaac 1080
 cccaaaccgg cttccacgc cgctcgtcaga ctgaagaccg aagactga 1128

<210> 272
 <211> 375
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(22)

<400> 272
 Met Phe Ile His Asn Ser Ile Cys Ser Ala Leu Cys Thr Ile Phe Leu
 1 5 10 15
 Ala Thr Ala Thr Met Gly Glu Asn Met Thr Leu Gln Glu Ala Phe Ala
 20 25 30
 Asp His Phe Tyr Val Gly Ala Ala Ile Ser Gln Arg Leu Phe Gln Pro
 35 40 45
 Asp Arg Ala Glu Thr Leu Gln Leu Ala Ala His Gln Phe Asn Ser Ile
 50 55 60
 Thr Ala Glu Asn Glu Met Lys Trp Gln Ser Leu Asn Pro Thr Pro Gly
 65 70 75 80
 Glu Tyr Arg Phe Glu Asn Ala Asp Lys Phe Val Arg Phe Gly Val Glu
 85 90 95
 Asn Asp Met Tyr Ile Val Gly His Val Leu Phe Trp His Ser Gln Thr
 100 105 110
 Pro Asp Trp Leu Phe Lys Asp Asp Gly Asn Phe Val Ser Arg Glu
 115 120 125
 Val Leu Leu Asp Arg Met Arg Ala His Val Arg Asn Leu Val Gln Arg
 130 135 140

Tyr Gly Asn His Val His Ala Trp Asp Val Ile Asn Glu Thr Phe Asn
 145 150 155 160
 Asp Asn Gly Ser Leu Arg Asp Ser Pro Trp Thr Gln Ile Leu Gly Glu
 165 170 175
 Glu Phe Ile Glu His Ala Phe Arg Ile Ala Gly Glu Glu Leu Pro Pro
 180 185 190
 His Val Glu Leu Leu Tyr Asn Asp Tyr Ser Met Thr Ile Pro Ala Lys
 195 200 205
 Arg Asp Ala Val Ala Glu Met Val Arg Asp Leu Ile Ala Lys Gly Ile
 210 215 220
 Arg Ile Asp Gly Val Gly Met Gln Gly His Trp Ala Arg Thr His Pro
 225 230 235 240
 Thr Ile Ala Asp Ile Glu Lys Ser Ile Leu Ala Phe Ala Gly Thr Gly
 245 250 255
 Val Gln Val His Ile Thr Glu Leu Asp Ile Asp Met Leu Pro Arg His
 260 265 270
 Pro Gln Met Phe Thr Gly Gly Ala Asp Thr Met Leu Arg Leu Gln Gln
 275 280 285
 Asp Pro Lys Leu Asp Pro Tyr Thr Glu Gly Leu Pro Ala Glu Asp Gln
 290 295 300
 Gln Ala Leu Ala Glu Arg Tyr Ala Ser Ile Phe Arg Leu Phe Leu Lys
 305 310 315 320
 His Ser Asp Val Ile Arg Arg Val Thr Phe Trp Gly Val Thr Asp Ala
 325 330 335
 His Thr Trp Leu Asn Asn Trp Pro Ile Arg Gly Arg Thr Ser His Pro
 340 345 350
 Leu Leu Phe Asp Arg Gln Asn Asn Pro Lys Pro Ala Phe His Ala Val
 355 360 365
 Val Arg Leu Lys Thr Glu Asp
 370 375

<210> 273

<211> 1134

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 273

atgggtttcat	cgctaatacaa	ttcttcatac	attcgggtca	agcactattc	gtgctcaagt	60
ttattgctcc	tgacattggc	agcctgtggc	ggccagcagc	ctcccccgga	tacgggatcc	120
agcacttcaa	gttcaagcag	ttcttcgagc	tccagttcaa	gcagctcttc	aagttccagc	180
tcaagcagtt	cttccagctc	cagctcgagc	agctcttcga	gttcgagctc	ttcatcatcc	240
agctcttcag	gggcaaacc	gccaccgacc	gggggcaagt	tcgtcggcaa	catcacgacc	300
cgaggcgccg	tccaagcgga	cttcattcag	tactgggatac	aaattacgcc	ggagaacgag	360
ggcaaatggg	gttctgtgga	aggaaactcg	gaccagtaca	actgggcgcc	tcttgatcgc	420
atctatgact	atgcacgtca	gcacaatatc	ccagtcaaag	cgcatagcgt	ggtttggggt	480
gcacaggtc	caggctggat	caacaatctg	agtcgpgccg	agcagcgatga	ggaaatcgag	540
gaatggattc	gtgattactg	cacgcgttac	ccagacaccc	aaatgatcga	cgtagttaac	600
gaggcgacc	caaatcacgc	ccccgctcgc	tatgcgagcga	atgccttcgg	caatgactgg	660
attaccgaag	cgttcaaact	ggcgcgccc	cactgcccga	acgccatttt	gatctacaac	720
gactataatt	tcatacttg	ggataccgat	gaaatcatgg	cgctgattcg	cccggctatc	780
gcagcagggg	tagtggatgc	ggtagggctg	caggcgcata	gcttgatcc	tgacgaatac	840
gctaacaaga	tgtggagtgc	cgctgaaata	cagcagaagc	tcgatctgat	ctctaccctt	900
ggcgtgccga	tgtatatttc	ggaatatgat	gtcgccaagt	ccaatgacca	agagcagttg	960
gcgattttca	gcgagcagtt	cccggctcct	tacgaacacc	ccaatgtcgt	aggtgtaacc	1020
ctctggggct	atattgatgg	agcgacgtgg	cgcgcgggct	cgggcttgat	tcgaaacggt	1080
cagcaccggc	ccgccatgca	atggctgctc	gagtacttgg	agaacaatcg	atag	1134

<210> 274

<211> 377

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(74)

<400> 274

```

Met Val Ser Ser Leu Ile Asn Ser Ser Tyr Ile Arg Leu Lys His Tyr
 1      5      10      15
Ser Cys Ser Ser Leu Leu Leu Thr Leu Ala Ala Cys Gly Gly Gln
 20      25      30
Gln Pro Pro Pro Asp Thr Gly Ser Ser Thr Ser Ser Ser Ser Ser
 35      40      45
Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
 50      55      60
Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
 65      70      75      80
Ser Ser Ser Gly Ala Asn Pro Pro Pro Thr Gly Gly Lys Phe Val Gly
 85      90      95
Asn Ile Thr Thr Arg Gly Ala Val Gln Ala Asp Phe Ile Gln Tyr Trp
100      105      110
Asp Gln Ile Thr Pro Glu Asn Glu Gly Lys Trp Gly Ser Val Glu Gly
115      120      125
Thr Arg Asp Gln Tyr Asn Trp Ala Pro Leu Asp Arg Ile Tyr Asp Tyr
130      135      140
Ala Arg Gln His Asn Ile Pro Val Lys Ala His Thr Leu Val Trp Gly
145      150      155      160
Ala Gln Ala Pro Gly Trp Ile Asn Asn Leu Ser Ala Ala Glu Gln Arg
165      170      175
Glu Glu Ile Glu Glu Trp Ile Arg Asp Tyr Cys Thr Arg Tyr Pro Asp
180      185      190
Thr Gln Met Ile Asp Val Val Asn Glu Ala His Pro Asn His Ala Pro
195      200      205
Ala Arg Tyr Ala Gln Asn Ala Phe Gly Asn Asp Trp Ile Thr Glu Ala
210      215      220
Phe Lys Leu Ala Arg Arg His Cys Pro Asn Ala Ile Leu Ile Tyr Asn
225      230      235      240
Asp Tyr Asn Phe Ile Thr Trp Asp Thr Asp Glu Ile Met Ala Leu Ile
245      250      255
Arg Pro Ala Ile Ala Ala Gly Val Val Asp Ala Val Gly Leu Gln Ala
260      265      270
His Ser Leu Tyr Pro Asp Glu Tyr Ala Asn Lys Met Trp Ser Ala Ala
275      280      285
Glu Ile Gln Gln Lys Leu Asp Leu Ile Ser Thr Leu Gly Val Pro Met
290      295      300
Tyr Ile Ser Glu Tyr Asp Val Ala Lys Ser Asn Asp Gln Glu Gln Leu
305      310      315      320
Ala Ile Phe Ser Glu Gln Phe Pro Val Leu Tyr Glu His Pro Asn Val
325      330      335
Val Gly Val Thr Leu Trp Gly Tyr Ile Asp Gly Ala Thr Trp Arg Ala
340      345      350
Gly Ser Gly Leu Ile Arg Asn Gly Gln His Arg Pro Ala Met Gln Trp
355      360      365
Leu Leu Glu Tyr Leu Glu Asn Asn Arg
370      375

```

<210> 275

<211> 1401

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 275

```

ttgggcgctg atccatttgc gctcacctat aacggaagag tgtacattta tatgtcgagt      60
gatgactatg aatatcacag caatggaacg attaaggata attcttttgc caatttgaat      120
agggtctttg tcatctcttc agcagatatg gtgaactgga cagatcatgg cgcgattcca      180
gtagctgggg caaatggcgc aaatggcggc aaaggaattg ccaaattgggc aggtgcttcc      240
tgggctccat cagcagcggg gaaaaaaatc aatgggaagg ataaattttt cctttatttc      300
gcgaacacgc gcggagggat tggcgttctg acagcagact ccccatcggg tccatggaca      360
gatacctatc gaaaagcact cgtcacgcca aatacaccag ggatggctgg agttgtatgg      420
ctttttgatc ctgccgtttt tgtagatgat gacggcactg gttatctata tgccggcgga      480

```

```

ggtgttccag gcggttctaa tccaacgcag ggacaatggg cgaatcctaa aacagcaaga 540
gttctaaaac taggacctga tatgacaagt gtggtaggca gcgcatcaac cattgatgct 600
ccttttatgt ttgaagattc ggggatgcat aagtataacg gaacctatta ctattcctat 660
tgcatcaact ttggcggtc ccacccagca gataaaccac ctgggtgagat cggttatatg 720
acgagctcaa gtccgatggg tccctttacg tatagagggc acttcctgaa aaatccgggt 780
gcatttttcg ggggagggcg taataacat catgctgtgt tcaattttta aaacgagtg 840
tatgtcgtgt atcataccca aacggtcagc tctgctttat acggatcagg aaaaggctac 900
agatctccgc atattaataa acttggtgcat aatgctgacg gctcccttcg agaagtcgca 960
gccaattttg aaggggttaa acagctttcc aacctgaatc cttatcagcg tgtagaagct 1020
gaaacattcg catggaatgg acgcattttg acagaggcat cttcagctcc aggcggaccg 1080
gtcaataacc agcatgtcac aaacattcaa aacggagatt ggggtggctgc cagtaacgct 1140
gatttcggat caaacggcg gagggacatt aaagcgaatg tagcatcaaa tacaggcggg 1200
aaaatagaag tacgcctcgg aagtcagac ggcagactcg tcggaacact gaatgtccct 1260
tccacagggg gaacaaataa ctggcgagaa gtagaaacgg cagtaaatgg agcagcaggc 1320
gtgcacaacg tattttttgt ttttactgga acaggtgcaa atctatttca atttgattcc 1380
tggcagttta ctcaaaggta a 1401

```

<210> 276

<211> 466

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 276

```

Met Gly Ala Asp Pro Phe Ala Leu Thr Tyr Asn Gly Arg Val Tyr Ile
1      5      10      15
Tyr Met Ser Ser Asp Asp Tyr Glu Tyr His Ser Asn Gly Thr Ile Lys
20      25      30
Asp Asn Ser Phe Ala Asn Leu Asn Arg Val Phe Val Ile Ser Ser Ala
35      40      45
Asp Met Val Asn Trp Thr Asp His Gly Ala Ile Pro Val Ala Gly Ala
50      55      60
Asn Gly Ala Asn Gly Gly Lys Gly Ile Ala Lys Trp Ala Gly Ala Ser
65      70      75      80
Trp Ala Pro Ser Ala Ala Val Lys Lys Ile Asn Gly Lys Asp Lys Phe
85      90      95
Phe Leu Tyr Phe Ala Asn Ser Gly Gly Ile Gly Val Leu Thr Ala
100      105      110
Asp Ser Pro Ile Gly Pro Trp Thr Asp Pro Ile Gly Lys Ala Leu Val
115      120      125
Thr Pro Asn Thr Pro Gly Met Ala Gly Val Val Trp Leu Phe Asp Pro
130      135      140
Ala Val Phe Val Asp Asp Asp Gly Thr Gly Tyr Leu Tyr Ala Gly Gly
145      150      155      160
Gly Val Pro Gly Gly Ser Asn Pro Thr Gln Gly Gln Trp Ala Asn Pro
165      170      175
Lys Thr Ala Arg Val Leu Lys Leu Gly Pro Asp Met Thr Ser Val Val
180      185      190
Gly Ser Ala Ser Thr Ile Asp Ala Pro Phe Met Phe Glu Asp Ser Gly
195      200      205
Met His Lys Tyr Asn Gly Thr Tyr Tyr Tyr Ser Tyr Cys Ile Asn Phe
210      215      220
Gly Gly Ser His Pro Ala Asp Lys Pro Pro Gly Glu Ile Gly Tyr Met
225      230      235      240
Thr Ser Ser Ser Pro Met Gly Pro Phe Thr Tyr Arg Gly His Phe Leu
245      250      255
Lys Asn Pro Gly Ala Phe Phe Gly Gly Gly Asn Asn His His Ala
260      265      270
Val Phe Asn Phe Lys Asn Glu Trp Tyr Val Val Tyr His Thr Gln Thr
275      280      285
Val Ser Ser Ala Leu Tyr Gly Ser Gly Lys Gly Tyr Arg Ser Pro His
290      295      300
Ile Asn Lys Leu Val His Asn Ala Asp Gly Ser Leu Arg Glu Val Ala
305      310      315      320
Ala Asn Phe Glu Gly Val Lys Gln Leu Ser Asn Leu Asn Pro Tyr Gln
325      330      335
Arg Val Glu Ala Glu Thr Phe Ala Trp Asn Gly Arg Ile Leu Thr Glu

```

340 345 350
 Ala Ser Ser Ala Pro Gly Gly Pro Val Asn Asn Gln His Val Thr Asn
 355 360 365
 Ile Gln Asn Gly Asp Trp Val Ala Ala Ser Asn Val Asp Phe Gly Ser
 370 375 380
 Asn Gly Ala Arg Thr Phe Lys Ala Asn Val Ala Ser Asn Thr Gly Gly
 385 390 395 400
 Lys Ile Glu Val Arg Leu Gly Ser Pro Asp Gly Arg Leu Val Gly Thr
 405 410 415
 Leu Asn Val Pro Ser Thr Gly Gly Thr Asn Asn Trp Arg Glu Val Glu
 420 425 430
 Thr Ala Val Asn Gly Ala Ala Gly Val His Asn Val Phe Phe Val Phe
 435 440 445
 Thr Gly Thr Gly Ala Asn Leu Phe Gln Phe Asp Ser Trp Gln Phe Thr
 450 455 460
 Gln Arg
 465

<210> 277
 <211> 1128
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 277
 atgcgaaca cattaatcct ttgattccg gccttgatga tgctttcttg cagtgcggga 60
 aatcaggata gaggaccatc cctgcacgcc gagttctcgg atgcattttt gattggaacg 120
 gcgctgaatt ctgagcagat attgggtcgg gatacacgcg gactcgaatt gattagaact 180
 cattttaacg ccattacgcc cgaaaacatt accaaatggg aggctatcca tcccgaaccc 240
 ggtgtctatg attttaaaga ggctgatgca ttcgtcgatt ttggccaaaa atataatatg 300
 ttcattggtg gtcatacact ggtttgccat agtcagacac cgcgctgggt cttcaaagac 360
 gaaaatggcg cgttggtatc gcgcgaggta ctgttagagc ggatgcgcga ccacatccac 420
 accgttggtg gccgctaccg tggacgtatt cacggctggg atgtcgtaaa cgaagccctc 480
 aatgaagacg gttcgtacag agaaacactg tggtagcaaa taattggtac ggactatatt 540
 cttaaagcat tcgaatttgc ccgggaggcc gatcccgacg ctgagctata ctataacgat 600
 tactcgcttg agaacccttc aaagagagcc ggcgcgatgc gaattgttca atacctgcag 660
 gaacatggtg ctccgattac tgggggttga acccaggggc atttcaccct cgactggccc 720
 gaactttctg aaattgaaca gaccgtcatt gattttgcct cccttggtat ggatgtaatt 780
 attaccgaat tggatatcga tgtactgcct cagccagacg attatactgg cgccgatgtg 840
 aatttttagc cagagcttta cgacgaactg aaccatggc ccaacggcct tccaccggaa 900
 attgaacagg aattggccaa tcgatatgcc gacatcttcg aaatctattt gcgtcatcgt 960
 gataaagtta cgcgagtgtg tttttgggtg gtcacagatg gcgactcgtg gaaaaataac 1020
 tggcctgtgc caggctgtac caactatccg ctcatctttg atcgaaactg gaagccaaaa 1080
 cccgcttttt tctcgaattgt tgatgcagcc agggaggcac tggattaa 1128

<210> 278
 <211> 375
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(19)

<400> 278
 Met Arg Asn Thr Leu Ile Leu Leu Ile Pro Ala Leu Met Met Leu Ser
 1 5 10 15
 Cys Ser Ala Gly Asn Gln Asp Arg Val Pro Ser Leu His Ala Glu Phe
 20 25 30
 Ser Asp Ala Phe Leu Ile Gly Thr Ala Leu Asn Ser Glu Gln Ile Leu
 35 40 45
 Gly Arg Asp Thr Arg Gly Leu Glu Leu Ile Arg Thr His Phe Asn Ala
 50 55 60
 Ile Thr Pro Glu Asn Ile Thr Lys Trp Glu Ala Ile His Pro Glu Pro
 65 70 75 80

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(22)

<400> 280

```

Met Leu Ser Pro Thr Arg Lys Leu Pro Pro Ala Ile Gly Leu Thr Phe
 1          5          10          15
Leu Phe Ala Ala Ser Ala Thr Pro Glu Thr Thr Leu Lys Asp Ala Phe
          20          25          30
Ala Asp His Phe Leu Val Gly Ala Leu Asn Glu Ser His Phe Ala
          35          40          45
Glu His Asn Pro Ala His Ala Gly Leu Val Ala Ala Asn Phe Asn Ala
          50          55          60
Ile Thr Ala Glu Asn Val Met Lys Trp Glu Ala Val His Pro Arg Pro
65          70          75          80
Gly Glu Tyr Thr Phe Gly Ala Ala Asp Arg Phe Val Glu Phe Gly Glu
          85          90          95
Lys Asn Gly Leu Phe Ile Val Gly His Thr Leu Ile Trp His Ser Gln
          100          105          110
Thr Pro Ala Trp Val Phe Glu Asp Glu Asn Gly Ala Pro Leu Gly Arg
          115          120          125
Glu Ala Leu Leu Glu Arg Met Arg Asp His Ile His Thr Val Ala Gly
          130          135          140
Arg Tyr Arg Gly Arg Val Lys Gly Trp Asp Val Val Asn Glu Ala Leu
145          150          155          160
Ala Glu Asp Gly Ser Leu Arg Asp Ser Pro Trp Arg Arg Ile Ile Gly
          165          170          175
Asp Asp Tyr Phe Val Lys Ala Phe Glu Phe Ala Arg Glu Ala Asp Pro
          180          185          190
Asp Ala Glu Leu Tyr Tyr Asn Asp Tyr Ser Ile Glu Asn Glu Pro Lys
          195          200          205
Arg Lys Gly Ala Val Ala Leu Val Arg Thr Leu Gln Ala Ala Gly Val
          210          215          220
Pro Val Ala Gly Val Gly Ile Gln Gly His Gly Asn Leu His Trp Pro
225          230          235          240
Ser Pro Arg Leu Val Glu Glu Ala Ile Arg Asp Phe Ala Ser Leu Gly
          245          250          255
Val Lys Val Met Ile
          260

```

<210> 281

<211> 963

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 281

```

gtgggcacct gcatgagcgg ggccgattcg cgcaaccctg cccgtctgga gctgatcaga      60
acgcagtaca gcatcatcac cccggaaaac gagctcaagc ccgattccgt tctggatgtg      120
gctgcccagcc gtgctctggc caaggaggac gataccgcg tggccgtgca tttcagcgcc      180
gccgctccca tcctgaactt tgcccgtgac aacggcatca aggtgcacgg tcatgtgctg      240
gtctggcaca gccagactcc cgaggagttc ttccacgagg gctataacgc ctccgcgccc      300
tatgtgagcc gcgaggtgat gctggcccgt ctggacaact acatccgtct catcttgaa      360
tatatggatg aaaactatcc cggcctgatc gtgtcctggg atgtggccaa cgaatgcgtg      420
gccgacggct ccaccgccct gcgcacctcc aactggacc gcgtggtggg gcaggatttt      480
gtggcccgcg ccttcgagat gcgccgataa tacgcgccg aagatgtgat gctctgtac      540
aacgattatt cactcccta tgagcccaag ctaccggca tcgtgaacct gctcaccgag      600
ctgacacagg agggtcatat cgacggctac ggcttcacga gccactacag tgtcggcgat      660
ccctccctgc aggcggtcga gaacgcgttc aaaaagatct ccgccctggg gctcaagctg      720
cgctgagcgc agctggacat caaggtagat gccgacagcg agcccaaccg cgcccttcag      780
gccgaccggt atgaggccct gctgcgcata tatatgaaat acggcgctcag cgccgtgcag      840
gtgtggggcg tatgcgacgg caccagctgg atcggcgca gctatcccct ccccttgac      900
gccgggctgc gtcccaagcc ctccttcttc ggcatactcc gcgcccttga cgaacagaac      960
tga                                         963

```

<210> 282

<211> 320
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 282
 Met Gly Thr Cys Met Ser Gly Ala Asp Ser Arg Asn Pro Ala Arg Leu
 1 5 10 15
 Glu Leu Ile Arg Thr Gln Tyr Ser Ile Ile Thr Pro Glu Asn Glu Leu
 20 25 30
 Lys Pro Asp Ser Val Leu Asp Val Ala Ala Ser Arg Ala Leu Ala Lys
 35 40 45
 Glu Asp Asp Thr Ala Val Ala Val His Phe Ser Ala Ala Ala Pro Ile
 50 55 60
 Leu Asn Phe Ala Arg Asp Asn Gly Ile Lys Val His Gly His Val Leu
 65 70 75 80
 Val Trp His Ser Gln Thr Pro Glu Glu Phe Phe His Glu Gly Tyr Asn
 85 90 95
 Ala Ser Ala Pro Tyr Val Ser Arg Glu Val Met Leu Ala Arg Leu Asp
 100 105 110
 Asn Tyr Ile Arg Leu Ile Phe Glu Tyr Met Asp Glu Asn Tyr Pro Gly
 115 120 125
 Leu Ile Val Ser Trp Asp Val Ala Asn Glu Cys Val Ala Asp Gly Ser
 130 135 140
 Thr Ala Leu Arg Thr Ser Asn Trp Thr Arg Val Val Gly Gln Asp Phe
 145 150 155 160
 Val Ala Arg Ala Phe Glu Ile Ala Asp Lys Tyr Ala Pro Glu Asp Val
 165 170 175
 Met Leu Cys Tyr Asn Asp Tyr Ser Thr Pro Tyr Glu Pro Lys Leu Thr
 180 185 190
 Gly Ile Val Asn Leu Leu Thr Glu Leu Thr Gln Glu Gly His Ile Asp
 195 200 205
 Gly Tyr Gly Phe Gln Ser His Tyr Ser Val Gly Asp Pro Ser Leu Gln
 210 215 220
 Ala Val Glu Asn Ala Phe Lys Lys Ile Ser Ala Leu Gly Leu Lys Leu
 225 230 235 240
 Arg Val Ser Glu Leu Asp Ile Lys Val Asp Ala Asp Ser Glu Pro Asn
 245 250 255
 Arg Ala Leu Gln Ala Asp Arg Tyr Glu Ala Leu Leu Arg Ile Tyr Met
 260 265 270
 Lys Tyr Gly Val Ser Ala Val Gln Val Trp Gly Val Cys Asp Gly Thr
 275 280 285
 Ser Trp Ile Gly Ala Ser Tyr Pro Leu Pro Phe Asp Ala Gly Leu Arg
 290 295 300
 Pro Lys Pro Ser Phe Phe Gly Ile Leu Arg Ala Leu Asp Glu Gln Asn
 305 310 315 320

<210> 283
 <211> 4161
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 283
 atgtataaaa gtttcgtcaa gaaagtctcc cttgtattat ctactctttt gctcttagtt 60
 tcggcgtttc ctgtctcata tgcacaaatg aattccatcc ccgtttatga agaaacgttt 120
 gaaaaccaag gaaactatgt ccaatctggg ggtgcgaccc tcactctagt aaaaaacaaa 180
 gtgtttgcag ggaatgaaga tggaactgca ctatatatta gtaatcgatc gaataactgg 240
 gacggggcag atttccgttt cacggatctt ggattacaag atggaaaaac atatacgatc 300
 aatattatag gatattgtcga tgaaaatgaa gttgttcctt caggagccca agtgtatttg 360
 caaactgtag ataaaacata tggatgggta gcaagcgagg acttaaaaaa cggagagtcg 420
 ttcactataa atacaacgtt cacccttgac atgagtaaag gggacaccgg tcttcgtata 480
 caatccaacg atagtggtaa aaaagtcca ttttacgtcg ggtatttttc aatttcaatt 540
 agtgatgtag aaggagaaga tgggtgggagc tctatttcaa ggccaccggc ttacctttt 600
 gaaactattg actttgaaga tcaaagtta agtggtattg agggacgagc aggcacggaa 660

acattgaccg	ttacgaatga	agcaaataga	actcctggag	gatcttatgc	actaaaagtg	720
gaaaaatagat	ctcaaaattg	gcatggacct	tccttacgca	tcgagaaata	tattgattta	780
ggttacgaat	atacaatttc	tctatgggtt	aaacttattt	caccacaag	tgacaaaatt	840
cagctttcta	cccaagtcgg	aagtggaaat	ggtgagagtt	ataacaatat	tttaagtaaa	900
gtaattagtg	ttgatgatgg	atgggtactg	tatgaaggaa	agtatcgcta	caatagttcg	960
ggaggggaat	atttaacaat	ttacgtagaa	agcccaaaaca	atagtactgc	atctttttac	1020
atcgatgata	ttcgtttaat	aaagagtggg	gacccaatct	ctgtacaaaa	agatctttctc	1080
cctatcaaga	gtgtttatga	agggtacttc	ttagtgggta	gtgccgtatc	agcgactgat	1140
ttagagggag	agagactcga	gcttctcaag	ttgcattaca	atagcataac	agcggaatac	1200
gccatgaaac	ctagctatit	acaacctact	aaaggaaact	ttaccttcga	agcagcagat	1260
agtattgtaa	ataaagccct	agaagaagga	atgaaagtac	atggacatgt	tctcgtatgg	1320
catcagcaga	cacctgaatg	gatgaccact	agagaagatg	gaagccctct	cggcagggaa	1380
gaagcgtag	aaaaatctaaa	aaatcacatt	gaaacagtta	tgaaacattt	tggtgataga	1440
gtaatttcat	gggatgttgt	caatgaagct	atcattgata	atccacctaa	tcctgataat	1500
tgggaggaat	cattaagaaa	atcaccatgg	tactattcaa	tcggttctga	ttatgttgag	1560
caagcatttc	gaattgcacg	acaagtittg	gacgaaaatg	gggtgggata	taagctatat	1620
tacaatgatt	acaattgaaga	taatcaaaga	aaagcacaag	ccatttacca	tatggtaaaa	1680
gagcttaatg	aaaaatatgc	acaagagcat	cctggtaaaa	gattaatcga	tggaattgga	1740
atgcaagggc	attacagtat	acgaacaaat	ccagataatg	tgaaaatgic	attagaaaga	1800
tttatttccc	ttggtgtgga	agttagtatt	actgaactcg	atattcaagc	tggaacggat	1860
aatcattcta	gggaagaaaca	gtcaaaaaga	caagcatatt	tatacgctaa	attattcaaa	1920
atattcaaa	aaaatgcatc	gcatatctcc	cgagttactg	tatggggatt	aaatgatgcg	1980
gcaagttggc	gtgcgtcaac	aagtccattg	ttatttgatc	gaaatttaca	ggccaaacca	2040
agttactatg	cggtaatgga	tcctgatata	tttatagaag	aaaatcctac	tgtgacagaa	2100
gagtcgcgga	aagcaattgc	tttgtatggt	atccctgtaa	ttgatggaag	catcgattcc	2160
atttgggaaa	gtgttcctta	catccctatt	gatcgttacc	aaatggcgtg	gcaaggagca	2220
agcggaaactg	ctaaagttct	ttgggacgaa	gggaatctgt	atgtattagt	acaagttaac	2280
gatgaccagt	tagataagtc	gagtacaaat	ccttgggagc	aggattcgat	agaggtgttt	2340
gtggacgaaa	ataatgcaaa	aacatcgatt	taccaagagg	atgatggaca	atatagagta	2400
aaatttgaca	acgaacatc	gtttaatcca	ccaagcattg	aaaatggatt	tatgtccgaa	2460
actaatgtat	ctgggactaa	ttatgtgggt	gagatgaaaa	tccctttaag	aagtatacaa	2520
ctaaaaaatg	ggcttgaaat	agggtttgat	gttcaaatga	atgatgggaa	aaatggtgct	2580
cgtcagagtg	tggctgcttg	gaatgatagc	actggaactg	catatatgga	tacatctgta	2640
ttcgggacac	ttactctttt	aaccacttta	gataatgaaa	atacaccagg	cagcggcaca	2700
acaccaggta	ttggcacaac	accaggcagt	ggcacaacac	caggcagcag	cacaacacca	2760
ggcagcggta	caacaccagg	tagcggcaca	acaccaggca	gtggcacaac	accaggcagc	2820
ggcacaacac	caggcagtg	cacaacacca	ggcagtgcca	caacaccggg	cagcggta	2880
acaccaggca	gtggcacaac	accaggcagt	ggtacaacac	cgggcagtg	cacaacacca	2940
gtgaagggtg	aaaatggtag	ggtgttttta	cagccgaaag	tagagacgaa	agaaaaagac	3000
ggcaaatgag	tagaaaaagt	ggcaactatt	tcaacaaatg	aagtigaagc	gattgtcaag	3060
gagctgtcga	atgaaaaata	acaagtcgtc	gtctccctcg	gtcgcgttcc	aaaagggtga	3120
gccacaaaag	tagatgtgcc	agctacatta	tttacacaag	cggaataata	acaagcagaa	3180
gcaacgattg	tgagcgccag	tgaacaagcg	acgtacaaat	tgccagtcga	agaagtgcag	3240
gcgtctcttg	cgacgattgc	ccggtcactc	ggtgcaacga	tagaacaagt	tagcatctcg	3300
attgaaatga	aagtgaacga	tcgcccgtca	ctacgtgtga	aaccgttggtc	tgatgcggta	3360
gagtttcatg	tcgtggcgaa	agcaaatgga	aaggaaacgc	tcacgcagtc	gtttactcaa	3420
tatgtcgaa	gcgaaatcgc	gttaaagcaa	tcggtcaacg	ctagtcgtgc	cattgcagtg	3480
cgcgtgaacg	atgacgggtc	acttacccta	gtaccgacaa	cgtttgttgg	caacaaagca	3540
gtcatataat	ctgtgacgaa	ctcgacgtat	gtttgttggg	aaggaaacaca	tacatttagt	3600
gacatccaac	cacattgggc	gaaagggtat	attgaaacac	tcgcccgaac	acagcttgct	3660
aaagggatga	cggacacaac	atatcgacca	aatgatcgga	tgacgcgcgc	gcaatttgct	3720
gtgttgctcg	tacgggcgct	aggattgccc	agcgaaacgt	atgacggtcg	ctttgctgat	3780
gtgaagggaa	cggagtgggt	taacaagaac	ggtgaattag	cagcggcagt	caagttcgga	3840
atcattcaag	gaaaaacagc	ttatatgttt	gcgcccgaatg	agccaatcac	tcgcccacaa	3900
gcagctgtca	tgatcgaaacg	ggcattgaaa	ctttcgatcg	ttggctatga	tgaggcaaca	3960
agcgacaaaa	cgaaaaaagt	gacagatttc	cgcgatgcaa	aacaattgcc	aacatgggca	4020
aaacaggcga	ttgaagcagt	ataccaagca	ggcatcatgc	aaggacgaga	tagcggaaac	4080
tttgatccga	caagccatgt	gacgcgtgcc	gaaatggcga	agggtgtaat	ggatatttta	4140
gagttgacaa	aacttattta	a				4161

<210> 284
 <211> 1386
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(28)

<400> 284

```

Met Tyr Lys Ser Phe Val Lys Lys Val Ser Leu Val Leu Ser Thr Leu
 1      5      10      15
Leu Leu Leu Val Ser Ala Phe Pro Val Ser Tyr Ala Gln Met Asn Ser
 20      25      30
Ile Pro Val Tyr Glu Glu Thr Phe Glu Asn Gln Gly Asn Tyr Val Gln
 35      40      45
Ser Gly Gly Ala Thr Leu Thr Leu Val Lys Asn Lys Val Phe Ala Gly
 50      55      60
Asn Glu Asp Gly Thr Ala Leu Tyr Ile Ser Asn Arg Ser Asn Asn Trp
 65      70      75      80
Asp Gly Ala Asp Phe Arg Phe Thr Asp Leu Gly Leu Gln Asp Gly Lys
 85      90      95
Thr Tyr Thr Ile Asn Ile Ile Gly Tyr Val Asp Glu Asn Glu Val Val
 100      105      110
Pro Ser Gly Ala Gln Val Tyr Leu Gln Thr Val Asp Lys Thr Tyr Gly
 115      120      125
Trp Leu Ala Ser Ala Asp Leu Lys Asn Gly Glu Ser Phe Thr Ile Asn
 130      135      140
Thr Thr Phe Thr Leu Asp Met Ser Lys Gly Asp Thr Arg Leu Arg Ile
 145      150      155      160
Gln Ser Asn Asp Ser Gly Lys Lys Val Ser Phe Tyr Val Gly Tyr Phe
 165      170      175
Ser Ile Ser Ile Ser Asp Val Glu Gly Asp Gly Gly Ser Ser Ile
 180      185      190
Ser Arg Pro Pro Ala Leu Pro Phe Glu Thr Ile Asp Phe Glu Asp Gln
 195      200      205
Ser Leu Ser Gly Phe Glu Gly Arg Ala Gly Thr Glu Thr Leu Thr Val
 210      215      220
Thr Asn Glu Ala Asn Arg Thr Pro Gly Gly Ser Tyr Ala Leu Lys Val
 225      230      235      240
Glu Asn Arg Ser Gln Asn Trp His Gly Pro Ser Leu Arg Ile Glu Lys
 245      250      255
Tyr Ile Asp Leu Gly Tyr Glu Tyr Thr Ile Ser Leu Trp Val Lys Leu
 260      265      270
Ile Ser Pro Thr Ser Ala Gln Ile Gln Leu Ser Thr Gln Val Gly Ser
 275      280      285
Gly Ser Gly Ala Ser Tyr Asn Ile Leu Ser Lys Val Ile Ser Val
 290      295      300
Asp Asp Gly Trp Val Leu Tyr Glu Gly Lys Tyr Arg Tyr Asn Ser Ser
 305      310      315      320
Gly Gly Glu Tyr Leu Thr Ile Tyr Val Glu Ser Pro Asn Asn Ser Thr
 325      330      335
Ala Ser Phe Tyr Ile Asp Asp Ile Arg Leu Ile Lys Ser Gly Asp Pro
 340      345      350
Ile Ser Val Gln Lys Asp Leu Leu Pro Ile Lys Ser Val Tyr Glu Gly
 355      360      365
Asp Phe Leu Val Gly Ser Ala Val Ser Ala Thr Asp Leu Glu Gly Glu
 370      375      380
Arg Leu Glu Leu Leu Lys Leu His Tyr Asn Ser Ile Thr Ala Glu Asn
 385      390      395      400
Ala Met Lys Pro Ser Tyr Leu Gln Pro Thr Lys Gly Asn Phe Thr Phe
 405      410      415
Glu Ala Ala Asp Ser Ile Val Asn Lys Ala Leu Glu Glu Gly Met Lys
 420      425      430
Val His Gly His Val Leu Val Trp His Gln Gln Thr Pro Glu Trp Met
 435      440      445
Thr Thr Arg Glu Asp Gly Ser Pro Leu Gly Arg Glu Glu Ala Leu Glu
 450      455      460
Asn Leu Lys Asn His Ile Glu Thr Val Met Lys His Phe Gly Asp Arg
 465      470      475      480
Val Ile Ser Trp Asp Val Val Asn Glu Ala Ile Ile Asp Asn Pro Pro
 485      490      495
Asn Pro Asp Asn Trp Glu Glu Ser Leu Arg Lys Ser Pro Trp Tyr Tyr
 500      505      510
Ser Ile Gly Ser Asp Tyr Val Glu Gln Ala Phe Arg Ile Ala Arg Gln
 515      520      525

```

Val⁵³⁰ Leu Asp Glu Asn Gly Trp⁵³⁵ Asp Ile Lys Leu Tyr⁵⁴⁰ Tyr Asn Asp Tyr
 Asn⁵⁴⁵ Glu Asp Asn Gln Arg⁵⁵⁰ Lys Ala Gln Ala Ile⁵⁵⁵ Tyr His Met Val Lys⁵⁶⁰
 Glu Leu Asn Glu Lys⁵⁶⁵ Tyr Ala Gln Glu His⁵⁷⁰ Pro Gly Lys Arg Leu Ile⁵⁷⁵
 Asp Gly Ile Gly⁵⁸⁰ Met Gln Gly His Tyr⁵⁸⁵ Ser Ile Arg Thr Asn Pro Asp⁵⁹⁰
 Asn Val Lys⁵⁹⁵ Met Ser Leu Glu Arg Phe Ile Ser Leu Gly Val Glu Val⁶⁰⁵
 Ser Ile Thr Glu Leu Asp Ile⁶¹⁰ Gln Ala Gly Thr Asp⁶²⁰ Asn His Leu Thr
 Glu⁶²⁵ Glu Gln Ser Lys Ala⁶³⁰ Gln Ala Tyr Leu Tyr⁶³⁵ Ala Lys Leu Phe Lys⁶⁴⁰
 Ile Phe Lys Glu Asn⁶⁴⁵ Ala Ser His Ile Ser⁶⁵⁰ Arg Val Thr Leu Trp Gly⁶⁵⁵
 Leu Asn Asp Ala⁶⁶⁰ Ala Ser Trp Arg Ala⁶⁶⁵ Ser Thr Ser Pro Leu Phe⁶⁷⁰
 Asp Arg Asn⁶⁷⁵ Leu Gln Ala Lys Pro⁶⁸⁰ Ser Tyr Tyr Ala Val Ile Asp Pro⁶⁸⁵
 Asp Thr Phe Ile Glu Glu Asn⁶⁹⁵ Pro Thr Val Thr Glu⁷⁰⁰ Glu Ser Arg Lys
 Ala Ile Ala Leu Tyr Gly⁷¹⁰ Ile Pro Val Ile Asp⁷¹⁵ Gly Ser Ile Asp Ser⁷²⁰
 Ile Trp Glu Ser Val⁷²⁵ Pro Tyr Ile Pro Ile⁷³⁰ Asp Arg Tyr Gln Met Ala⁷³⁵
 Trp Gln Gly Ala⁷⁴⁰ Ser Gly Thr Ala Lys⁷⁴⁵ Val Leu Trp Asp Glu Gly Asn⁷⁵⁰
 Leu Tyr Val⁷⁵⁵ Leu Val Gln Val Asn⁷⁶⁰ Asp Asp Gln Leu Asp Lys Ser Ser⁷⁶⁵
 Thr Asn Pro Trp Glu Gln Asp⁷⁷⁵ Ser Ile Glu Val Phe⁷⁸⁰ Val Asp Glu Asn
 Asn Ala Lys Thr Ser Phe Tyr⁷⁹⁰ Gln Glu Asp Asp⁷⁹⁵ Gly Gln Tyr Arg Val⁸⁰⁰
 Asn Phe Asp Asn Glu⁸⁰⁵ Thr Ser Phe Asn Pro⁸¹⁰ Pro Ser Ile Glu Asn Gly⁸¹⁵
 Phe Met Ser Glu⁸²⁰ Thr Asn Val Ser Gly⁸²⁵ Thr Asn Tyr Val Val Glu Met⁸³⁰
 Lys Ile Pro⁸³⁵ Leu Arg Ser Ile Gln⁸⁴⁰ Leu Lys Asn Gly Ser Glu Ile Gly⁸⁴⁵
 Phe Asp Val Gln Ile Asn Asp⁸⁵⁵ Gly Lys Asn Gly Ala Arg Gln Ser Val⁸⁶⁰
 Ala Ala Trp Asn Asp Thr⁸⁷⁰ Gly Thr Ala Tyr⁸⁷⁵ Met Asp Thr Ser Val⁸⁸⁰
 Phe Gly Thr Leu Thr⁸⁸⁵ Leu Thr Thr Leu Asp⁸⁹⁰ Asn Glu Asn Thr Pro⁸⁹⁵
 Gly Ser Gly Thr⁹⁰⁰ Thr Pro Gly Ser Gly⁹⁰⁵ Thr Thr Pro Gly Ser Gly Thr⁹¹⁰
 Thr Pro Gly⁹¹⁵ Ser Ser Thr Thr Pro⁹²⁰ Gly Ser Gly Thr Thr Pro Gly Ser⁹²⁵
 Gly Thr Thr Pro Gly Ser Gly⁹³⁵ Thr Thr Pro Gly Ser Gly Thr Thr Pro⁹⁴⁰
 Gly Ser Gly Thr Thr Pro⁹⁵⁰ Gly Ser Gly Thr Thr Pro Gly Ser Gly Thr⁹⁶⁰
 Thr Pro Gly Ser Gly⁹⁶⁵ Thr Thr Pro Gly Ser Gly Thr Thr Pro Gly Ser⁹⁷⁵
 Gly Thr Thr Pro Val Lys Gly Glu Asn⁹⁸⁵ Gly Thr Val Val Leu Gln Pro⁹⁹⁰
 Lys Val Glu Thr Lys Glu Lys Asp¹⁰⁰⁰ Gly Lys Val Val Glu Lys Val Ala¹⁰⁰⁵
 Thr Ile Ser Thr Asn Glu Val¹⁰¹⁵ Glu Ala Ile Val Lys¹⁰²⁰ Glu Leu Ser Asn
 Glu¹⁰²⁵ Asn Lys Gln Val Val¹⁰³⁰ Ser Leu Gly Ser Leu Pro Lys Gly Val¹⁰⁴⁰
 Ala Thr Lys Val Asp Val Pro Ala Thr Leu¹⁰⁵⁰ Phe Thr Gln Ala Ala Asn¹⁰⁵⁵
 Lys Gln Ala Glu Ala Thr Ile Val Ser Ala Ser Glu Gln Ala Thr Tyr¹⁰⁷⁰
 Lys Leu Pro Val Lys Glu Val Gln Ala Ser Leu Ala Thr Ile Ala Arg

1075 1080 1085
 Ser Leu Gly Ala Thr Ile Glu Gln Val Ser Ile Ser Ile Glu Met Lys
 1090 1095 1100
 Val Asn Asp Ala Pro Ser Leu Arg Val Lys Pro Leu Ser Asp Ala Val
 1105 1110 1115 1120
 Glu Phe His Val Val Ala Lys Ala Asn Gly Lys Glu Arg Val Ile Asp
 1125 1130 1135
 Arg Phe Thr Gln Tyr Val Glu Arg Glu Ile Ala Leu Lys Gln Ser Val
 1140 1145 1150
 Asn Ala Ser Arg Ala Ile Ala Val Arg Val Asn Asp Asp Gly Ser Leu
 1155 1160 1165
 Thr Pro Val Pro Thr Thr Phe Val Gly Asn Lys Ala Val Ile Lys Ser
 1170 1175 1180
 Leu Thr Asn Ser Thr Tyr Val Val Val Glu Gly Thr His Thr Phe Ser
 1185 1190 1195 1200
 Asp Ile Gln Pro His Trp Ala Lys Gly Tyr Ile Glu Thr Leu Ala Ala
 1205 1210 1215
 Lys Gln Leu Val Lys Gly Met Thr Asp Thr Thr Tyr Arg Pro Asn Asp
 1220 1225 1230
 Arg Met Thr Arg Ala Gln Phe Ala Val Leu Leu Val Arg Ala Leu Gly
 1235 1240 1245
 Leu Pro Ser Glu Thr Tyr Asp Gly Arg Phe Ala Asp Val Lys Gly Thr
 1250 1255 1260
 Glu Trp Phe Asn Lys Asn Gly Glu Leu Ala Ala Val Lys Phe Gly
 1265 1270 1275 1280
 Ile Ile Gln Gly Lys Thr Ala Tyr Met Phe Ala Pro Asn Glu Pro Ile
 1285 1290 1295
 Thr Arg Ala Gln Ala Ala Val Met Ile Glu Arg Ala Leu Lys Leu Ser
 1300 1305 1310
 Ile Val Gly Tyr Asp Glu Ala Thr Ser Asp Lys Thr Lys Lys Val Thr
 1315 1320 1325
 Asp Phe Arg Asp Ala Lys Gln Leu Pro Thr Trp Ala Lys Gln Ala Ile
 1330 1335 1340
 Glu Ala Val Tyr Gln Ala Gly Ile Met Gln Gly Arg Asp Ser Gly Asn
 1345 1350 1355 1360
 Phe Asp Pro Thr Ser His Val Thr Arg Ala Glu Met Ala Lys Val Leu
 1365 1370 1375
 Met Asp Ile Leu Glu Leu Thr Lys Leu Ile
 1380 1385

<210> 285
 <211> 1569
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 285
 gtgaaccgaa acacatctcc ggacttcaag cttggagcag gcgttacggt aagtgatttc 60
 ctgaacaagg gcaggcaata tgaggtcgct gtggcgaatc ttgatgaaat gacggccgga 120
 aatgccatga agcagtcttc tgtaatgagg cctgacggat ccatggattt cactcagggtc 180
 agaagattca tcgaggaggc cgaacgtgtc ggaatgacag tgtacggcca tacattggca 240
 tggcattcac agcagcagaa cgcctatctt aacggctctga tcaagggcaa gaagaccgag 300
 gtcgagccag gccaggagtc agaggtcggt cttctccaga cagatttcaa tgacggaaat 360
 gtcacattca acggatgggg aaacaattct tcaaggactg tcgagaatgg tgcattaaag 420
 cttacaaacc cttctgtagt aaacagttgg gagggcccagt tcgcatatga tttttcagag 480
 gccttcgaga tggacaagac atataagctc aagttcagga tcaagggctc ggctgcagga 540
 aagatcgagg caggcttcca gatcactgac ggctaccttt cggcagggtga gttcggaaac 600
 gtagagttca ataccagtg gaaggatgtc gagctctcat gcgtatgttc cgctgaaggc 660
 ggtacacgct tgatcttcag tttcggagag ttgcccggag atatctatat cgacgatttc 720
 tgcttcagtg tgggaaggagc tggatatatc tacgaagatc ttaccccggc agagaagaag 780
 gagcgcctta ccgaggcaat ggaccgttgg atcaagggaa tgatggagggt taccgctacc 840
 aggggtttctg cgtgggatgc tgtcaatgag gcgatttccg gccgtgatac aaatggcgac 900
 ggcttctatg aacttgagtc ggcacaatgg ggaagctcga acaacttcta ttggcaggat 960
 tatctcggct caggagatta tgtgcgtatc gtgatcgcaa agggccgcaa gtattatgag 1020
 gaattcggcg atcaggctcc tttgagactc ttcatcaatg actacaacct cgaatctgac 1080
 tgggatgaca acaagaagct caagagcctt atccattgga tcggtgtctg ggagtctgac 1140
 ggagtgacaa agatcgacgg aatcggtacc cagatgcacg ttctgtatta cgagaatcct 1200

gatattcagg	caagcaagga	gaaacattat	gtgcagatgc	ttcagcttat	ggcaaataca	1260
ggaaagctcg	tgaagatctc	cgagcttgat	atgggctatg	tagaccgcaa	cgaaataact	1320
gtgggaacag	cggacatgac	cgaccagcag	catagggcca	tggcggatta	ttatgacttc	1380
atcgtgcgca	agtactttga	gatcgtgcct	cctgcacagc	agtatggcat	cacgcagtgg	1440
tgcattgacgg	atgctcccg	agctatcggc	acaggctgga	gaggcgggtga	gcctgtgggc	1500
ctgtgggacc	agaattacaa	ccgcaagtat	gcgtacgcag	gatttgcaaa	cggacttaga	1560
gcgaaataa						1569

<210> 286

<211> 522

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 286

Met	Asn	Arg	Asn	Thr	Ser	Pro	Asp	Phe	Lys	Leu	Gly	Ala	Gly	Val	Thr
1				5					10					15	
Val	Ser	Asp	Phe	Leu	Asn	Lys	Gly	Arg	Gln	Tyr	Glu	Val	Ala	Val	Ala
			20					25					30		
Asn	Leu	Asp	Glu	Met	Thr	Ala	Gly	Asn	Ala	Met	Lys	Gln	Ser	Ser	Val
		35					40					45			
Met	Arg	Pro	Asp	Gly	Ser	Met	Asp	Phe	Thr	Gln	Val	Arg	Arg	Phe	Ile
	50					55					60				
Glu	Glu	Ala	Glu	Arg	Val	Gly	Met	Thr	Val	Tyr	Gly	His	Thr	Leu	Ala
65					70					75				80	
Trp	His	Ser	Gln	Gln	Gln	Asn	Ala	Tyr	Leu	Asn	Gly	Leu	Ile	Lys	Gly
			85						90					95	
Lys	Lys	Thr	Glu	Val	Glu	Pro	Gly	Gln	Glu	Ser	Glu	Val	Val	Leu	Leu
			100					105					110		
Gln	Thr	Asp	Phe	Asn	Asp	Gly	Asn	Val	Thr	Phe	Asn	Gly	Trp	Gly	Asn
		115					120					125			
Asn	Ser	Ser	Arg	Thr	Val	Glu	Asn	Gly	Ala	Leu	Lys	Leu	Thr	Asn	Pro
		130				135					140				
Ser	Val	Val	Asn	Ser	Trp	Glu	Ala	Gln	Phe	Ala	Tyr	Asp	Phe	Ser	Glu
145					150					155					160
Ala	Phe	Glu	Met	Asp	Lys	Thr	Tyr	Lys	Leu	Lys	Phe	Arg	Ile	Lys	Gly
			165						170					175	
Ser	Ala	Ala	Gly	Lys	Ile	Ala	Ala	Gly	Phe	Gln	Ile	Thr	Asp	Gly	Tyr
			180					185					190		
Leu	Ser	Ala	Gly	Glu	Phe	Gly	Thr	Val	Glu	Phe	Asn	Thr	Gln	Trp	Lys
		195					200					205			
Asp	Val	Glu	Leu	Ser	Cys	Val	Cys	Ser	Ala	Glu	Gly	Gly	Thr	Arg	Leu
		210				215					220				
Ile	Phe	Ser	Phe	Gly	Glu	Phe	Ala	Gly	Asp	Ile	Tyr	Ile	Asp	Asp	Phe
225					230					235					240
Cys	Phe	Ser	Val	Glu	Gly	Ala	Gly	Tyr	Ile	Tyr	Glu	Asp	Leu	Thr	Pro
			245						250					255	
Ala	Glu	Lys	Lys	Glu	Arg	Leu	Thr	Glu	Ala	Met	Asp	Arg	Trp	Ile	Lys
			260					265					270		
Gly	Met	Met	Glu	Val	Thr	Ala	Thr	Arg	Val	Ser	Ala	Trp	Asp	Ala	Val
		275					280					285			
Asn	Glu	Ala	Ile	Ser	Gly	Arg	Asp	Thr	Asn	Gly	Asp	Gly	Phe	Tyr	Glu
		290				295					300				
Leu	Glu	Ser	Ala	Gln	Trp	Gly	Ser	Ser	Asn	Asn	Phe	Tyr	Trp	Gln	Asp
305					310					315					320
Tyr	Leu	Gly	Ser	Gly	Asp	Tyr	Val	Arg	Ile	Val	Ile	Ala	Lys	Ala	Arg
			325						330					335	
Lys	Tyr	Tyr	Glu	Glu	Phe	Gly	Gly	Thr	Ala	Pro	Leu	Arg	Leu	Phe	Ile
			340					345					350		
Asn	Asp	Tyr	Asn	Leu	Glu	Ser	Asp	Trp	Asp	Asp	Asn	Lys	Lys	Leu	Lys
		355					360					365			
Ser	Leu	Ile	His	Trp	Ile	Gly	Val	Trp	Glu	Ser	Asp	Gly	Val	Thr	Lys
		370				375					380				
Ile	Asp	Gly	Ile	Gly	Thr	Gln	Met	His	Val	Ser	Tyr	Tyr	Glu	Asn	Pro
385					390					395					400
Asp	Ile	Gln	Ala	Ser	Lys	Glu	Lys	His	Tyr	Val	Gln	Met	Leu	Gln	Leu
			405						410					415	

Met Ala Asn Thr Gly Lys Leu Val Lys Ile Ser Glu Leu Asp Met Gly
 Tyr Val Asp Arg Asn Gly Asn Thr Val Gly Thr Ala Asp Met Thr Asp
 Gln Gln His Arg Ala Met Ala Asp Tyr Tyr Asp Phe Ile Val Arg Lys
 Tyr Phe Glu Ile Val Pro Pro Ala Gln Gln Tyr Gly Ile Thr Gln Trp
 Cys Met Thr Asp Ala Pro Gly Ala Ile Gly Thr Gly Trp Arg Gly Gly
 Glu Pro Val Gly Leu Trp Asp Gln Asn Tyr Asn Arg Lys Tyr Ala Tyr
 Ala Gly Phe Ala Asn Gly Leu Arg Ala Lys

<210> 287

<211> 1695

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 287

atgactattc	atctacaaaa	aaatctgctc	tttgcagcag	tcttactgac	agggcaagca	60
gcgcattg	tcacctccgg	atctggcgag	gccacgctgg	atgttaataa	cagctggggg	120
tcagggtatt	gcgctaaccgt	caccattgcc	aacaacggca	gccaagccat	tacttcctgg	180
acggttggcc	tgaaccttaa	cggcacgacg	atcaataatc	tttggaaagg	taatttgagt	240
gggggttacg	ttacccccgt	ggcctacaat	gcgaatgtgg	cgccgggtgc	taataccagc	300
tttggtttct	gtgctaattg	cagcgcaacc	cccacgttgg	cgacgtttga	agttcaaggc	360
ggcggtggcg	cagccagtag	ttccagcagt	tctgctattt	caagcagctc	cagcagttcg	420
gttgccggcg	gcagcaacag	cgtaaccgta	cgcatgagcg	gcgtaaccgg	agacgaaagc	480
gtgagcctgg	aaatcggtgg	tcagaccatc	gagacctgga	ccctgagcgt	cggtatgctc	540
gactacaccg	tacagaccaa	cgctaccggg	gagttgctgt	ttgcgtttac	caatgatgaa	600
gggtaccgcg	atgtcgaagt	ggattacatc	atggtcaacg	gcgtgacctc	tcaggccgaa	660
gatcaggaag	ataacaccgg	cgcggtggac	ggagagtgtg	gtgcgggttc	cttctcgagc	720
atgctgcact	gcaacggttc	tatcggtctt	ggcaatccat	tcgatggcag	caactcgtcg	780
tccagctcca	gcaacacttc	gagctcgagc	agcagctcta	accctaacc	tggcaatccg	840
aatttccccg	atttcttcgt	aggtaacatc	accaccagcg	gctcagttcg	ctctgacttc	900
atgcagtatt	gggatcaaat	tacgccggaa	aacgaaggta	aatgggggtc	cgtagaggga	960
actcgcgacc	agtacaactg	gggtccgctg	gatgcaattt	acaatttcgc	ccgtgcgaac	1020
gggtattcctg	taaaagcgca	cactctggtg	tggggcagtc	agcaaccggg	ttggattggc	1080
ggcttgagtg	ccgctgagca	gcgcgcggaa	attgaagagt	ggattcgcga	ttactgcgca	1140
cgctaccggg	atacggcgat	gatcgatgtg	gtgaatgaag	ctctgccttc	gcacgctccg	1200
gcaaaactatg	cgcccaatgc	atttggaaagc	gattggatca	ctgaatcctt	ccgtttggca	1260
cgctcagtact	gtccggatgc	cggtgctgac	tacaacgact	ataacttcac	gacctgggac	1320
accgatgcca	tcattcagat	gattcgccca	gccgtgaact	ccggttatgt	cgatgacttg	1380
ggctttcaag	gcacagcctt	gtattcccca	caagtctgga	ccgctcagca	aatccagagc	1440
aaatttgatc	agatttccga	gttgggcttg	ccgtgtaca	tctccgagta	cgacatcgaa	1500
gcaacgaatg	atcagactca	gttacagtac	atgcagatgc	acttcccgat	cttctacaac	1560
catcccaacg	tggccgggtat	taccctttgg	ggttatgttg	tgggtgctac	ctggcgatgat	1620
ggcactggcc	tgatccaaag	caatgggtcag	cagcgcccgg	ccatgcagtg	gttgatggag	1680
tatctaaacc	gctaa					1695

<210> 288

<211> 564

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(23)

<400> 288

Met Thr Ile His Leu Gln Lys Asn Leu Leu Phe Ala Ala Val Leu Leu
 1 5 10 15
 Thr Gly Gln Ala Ala His Ala Leu Thr Ser Gly Ser Gly Glu Ala Thr

Page 211

<210> 289
 <211> 2796
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 289
 atgaagttca ctttgacacc gctgctgtgc gggttcgcct tattgttggg ttgcgcggtg 60
 caggcaaccc cagccgcttc gttaaagcag gcctatcagc cgtttttcca tatcggcacc 120
 gcagtcagtc tggcgcaatt acaaccatcc aaagaacatg aacgcgcttt aattgcgag 180
 cactttaaca gtctgaccgc cgaaaacctg atgaaatggg aggaattca acccacggaa 240
 ggcaactttg attttaaagc ggccgatcag ttggttgctg ttgccgaaca acatcaaag 300
 tggatgatcg gccataccat tctgtggcat gaacaaaccc cagactgggt gtttcagggg 360
 ctggatggca aaccggccag caagcagctg ctactggccc gcttgacca acatatccaa 420
 acggtcgttg gccgttacca gggccgggtc aatggctggg atgtggtgaa tgaagcgctc 480
 aatgaagatg gcagcctgcg cgataccccc tggcggcgca ttttgggtga tgattacatt 540
 gccaccactt ttgctgtggg gcatcaggtc gaccctaaag ccaaactcta ttacaacgat 600
 tacaacctgt ttaaaccgga aaaacgcgcc ggggtgctgc ggattatcca acaactgcag 660
 caaaaaaatg tgccatttga tgccattggt gaacaagcgc attacggcct tgattcgctc 720
 gcattcaaag acgttgaaga ttcatcaac gcttttgctg ccaccggcct ggacgtgatg 780
 ctaaccgaac tggagatttc agtattgctg tatccatccg gcatgacgca ggggtgccgat 840
 atcagtcagc atcaggaatt gcaggaacaa ctaaaccctt atcgcgatgg tttgcccaa 900
 gccgtcgaac aagcctggca acaacgggat ctgctattgt tttcgtgtt attacgccag 960
 cagcaaaaac tgcattcgct gaccttctgg ggccttagat atggccaaag ctggcgtaat 1020
 aatttcccga tgcgcggctg caccgattac ccactgctgt ttgatcgcaa gctgcaagcc 1080
 aaaccgctgt taagcgcaat gacggcatta gccgcagacc agactaaagc caagcccaa 1140
 atgaatcagc tgggctttgc gccgacttcg accaaactgt tgattgtgca gggtcggcaa 1200
 tcagtgcctt ttcattgttt ggataccgag accggccaaa cgggtgctgca aggccaaagt 1260
 tcggcgccca ggttttgccc tgaatcgggg gaatgggtca gtgctgccga tttttctgca 1320
 gtgataactc ccggcaccta tcagatcaac atctcaggaa cgccgccaca aactgtcaag 1380
 atccagcccg aaccctatgc cgcgtgcat gatgcggcaa tcaaaagccta ttattttaac 1440
 cgcgcctcgc tcacactgga gccaaagtgt gccggacctt gggcacgcgc agcggggcat 1500
 ccggatacca aagtaggggt gcatgcttct gctgcatcgg gctgcatcgg agaaggttat 1560
 gagctcagcg ctgcaaaagg ctggtatgac gccggtgact acaacaaata cgtggtgaat 1620
 tccggcatta ccagttacac cctgttgtag gcctggcagg attttcctga gttttatcaa 1680
 agccggacct ggaatattcc ggagtccggc aacgcggtac cggacattct cgacgaaacc 1740
 ttatggaatc tgcagtgggt cagcgccatg caagacccaa acgacggggg cgtctatcac 1800
 aagctgactg aactgaattt ttccggcaac caaatgccgg accaagtga agcagagcgt 1860
 tatgtggtgc aaaaaaccac cgccgcggca ctgaatttcg ctgcggtgtt ggccaaagcc 1920
 agtacggttt ttgcaaaatt tgacgccag ttgcccggcc tgcgcaaca ataccgtcag 1980
 caagcactgc tcgcctggca atgggcgcaa aaaaatccgc agcaaacta tcaacaaccc 2040
 aaagatgtcc acactggcgc ttatgtgtag aaacaactgg ctgatgaatg ggcctgggct 2100
 ggcgcggagc tgtattttat gaccggcgag caaagtatat tgcagccact gttggcgctg 2160
 gacacgccaa tcagtgcagc atcctgggcc agtgtcagcg ctttggggta tttttcttg 2220
 gcttcggcga aacagcttga gcccgcacta cggcaacagg tacaacagaa aatccaacaa 2280
 gccgcgcgcg aaatcctgca ggaacatcaa acatccgcct atcaggtggc gatgaccaa 2340
 aacgattttg tctggggcag taatgcggtg gcaatgaata aagcgatgtt gttataccag 2400
 gcgtggaaaa tagcgccaaa accggagctg ctacaggcga tgcaaggctt ggttgattac 2460
 gttttggggc gcaaccggtt gcagcagctt tatgtcacag ggtttggcga gcaaagccc 2520
 cagcagatcc accaccgacc ttccggccgc gatgccatca aagcgcgggt accaggttgg 2580
 ttagtcggtg gtgcacagcc gggtaagcag gataaatgca cttatgccgg cgctttacc 2640
 gctgtcggcg ctttacccgc tgccagcacc ttaccagcca ccacttatct tgatgactgg 2700
 tgcagttacg ccaccaacga agtggcgatt aactggaatg cacctttggt gtatgtgctg 2760
 gcatggcacc tttcgcaaaa caccaagaca ccaata 2796

<210> 290
 <211> 931
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(22)

<400> 290
 Met Lys Phe Thr Leu Thr Pro Leu Leu Cys Gly Phe Ala Leu Leu Leu
 1 5 10 15
 Gly Cys Ala Val Gln Ala Thr Pro Ala Ala Ser Leu Lys Gln Ala Tyr
 20 25 30
 Gln Pro Phe Phe His Ile Gly Thr Ala Val Ser Leu Ala Gln Leu Gln
 35 40 45
 Pro Ser Lys Glu His Glu Arg Ala Leu Ile Ala Gln His Phe Asn Ser
 50 55 60
 Leu Thr Ala Glu Asn Leu Met Lys Trp Glu Glu Ile Gln Pro Thr Glu
 65 70 75 80
 Gly Asn Phe Asp Phe Lys Ala Ala Asp Gln Leu Val Ala Phe Ala Glu
 85 90 95
 Gln His Gln Met Trp Met Ile Gly His Thr Ile Leu Trp His Glu Gln
 100 105 110
 Thr Pro Asp Trp Val Phe Gln Gly Leu Asp Gly Lys Pro Ala Ser Lys
 115 120 125
 Gln Leu Leu Leu Ala Arg Leu Thr Lys His Ile Gln Thr Val Val Gly
 130 135 140
 Arg Tyr Gln Gly Arg Val Asn Gly Trp Asp Val Val Asn Glu Ala Leu
 145 150 155 160
 Asn Glu Asp Gly Ser Leu Arg Asp Thr Pro Trp Arg Arg Ile Leu Gly
 165 170 175
 Asp Asp Tyr Ile Ala Thr Thr Phe Ala Leu Val His Gln Val Asp Pro
 180 185 190
 Lys Ala Lys Leu Tyr Tyr Asn Asp Tyr Asn Leu Phe Lys Pro Glu Lys
 195 200 205
 Arg Ala Gly Val Leu Arg Ile Ile Gln Gln Leu Gln Gln Lys Asn Val
 210 215 220
 Pro Ile His Ala Ile Gly Glu Gln Ala His Tyr Gly Leu Asp Ser Pro
 225 230 235 240
 Ala Phe Lys Asp Val Glu Asp Ser Ile Asn Ala Phe Ala Ala Thr Gly
 245 250 255
 Leu Asp Val Met Leu Thr Glu Leu Glu Ile Ser Val Leu Pro Tyr Pro
 260 265 270
 Ser Gly Met Thr Gln Gly Ala Asp Ile Ser Gln His Gln Glu Leu Gln
 275 280 285
 Glu Gln Leu Asn Pro Tyr Arg Asp Gly Leu Pro Lys Ala Val Glu Gln
 290 295 300
 Ala Trp Gln Gln Arg Tyr Leu Asp Leu Phe Ser Leu Leu Arg Gln
 305 310 315 320
 Gln Gln Lys Leu His Arg Val Thr Phe Trp Gly Leu Asp Asp Gly Gln
 325 330 335
 Ser Trp Arg Asn Asn Phe Pro Met Arg Gly Arg Thr Asp Tyr Pro Leu
 340 345 350
 Leu Phe Asp Arg Lys Leu Gln Ala Lys Pro Leu Leu Ser Ala Leu Thr
 355 360 365
 Ala Leu Ala Ala Asp Gln Thr Lys Ala Lys Pro Lys Met Asn Gln Leu
 370 375 380
 Gly Phe Ala Pro Thr Ser Thr Lys Leu Leu Ile Val Pro Gly Arg Gln
 385 390 395 400
 Ser Val Pro Phe His Val Leu Asp Thr Glu Thr Gly Gln Thr Val Leu
 405 410 415
 Gln Gly Gln Ser Ser Ala Ala Arg Phe Trp Pro Glu Ser Gly Glu Trp
 420 425 430
 Val Ser Ala Ala Asp Phe Ser Ala Val Ile Thr Pro Gly Thr Tyr Gln
 435 440 445
 Ile Asn Ile Ser Gly Thr Pro Gln Thr Val Lys Ile Gln Ala Glu
 450 455 460
 Pro Tyr Ala Ala Leu His Asp Ala Ala Ile Lys Ala Tyr Tyr Phe Asn
 465 470 475 480
 Arg Ala Ser Leu Thr Leu Glu Pro Lys Phe Ala Gly Pro Trp Ala Arg
 485 490 495
 Ala Ala Gly His Pro Asp Thr Lys Val Arg Val His Ala Ser Ala Ala
 500 505 510
 Ser Ala Ser Arg Pro Glu Gly Tyr Glu Leu Ser Ala Ala Lys Gly Trp
 515 520 525
 Tyr Asp Ala Gly Asp Tyr Asn Lys Tyr Val Val Asn Ser Gly Ile Thr
 530 535 540

Ser Tyr Thr Leu Leu Gln Ala Trp Gln Asp Phe Pro Glu Phe Tyr Gln
 545 550
 Ser Arg Thr Trp Asn Ile Pro Glu Ser Gly Asn Ala Val Pro Asp Ile
 565 570 575
 Leu Asp Glu Thr Leu Trp Asn Leu Gln Trp Phe Ser Ala Met Gln Asp
 580 585 590
 Pro Asn Asp Gly Gly Val Tyr His Lys Leu Thr Glu Leu Asn Phe Ser
 595 600 605
 Ala Thr Gln Met Pro Asp Gln Val Thr Ala Glu Arg Tyr Val Val Gln
 610 615 620
 Lys Thr Thr Ala Ala Ala Leu Asn Phe Ala Ala Val Leu Ala Lys Ala
 625 630 635 640
 Ser Thr Val Phe Ala Lys Phe Asp Ala Gln Leu Pro Gly Leu Ser Gln
 645 650 655
 Gln Tyr Arg Gln Gln Ala Leu Leu Ala Trp Gln Trp Ala Gln Lys Asn
 660 665 670
 Pro Gln Gln Ile Tyr Gln Gln Pro Lys Asp Val His Thr Gly Ala Tyr
 675 680 685
 Gly Asp Lys Gln Leu Ala Asp Glu Trp Ala Trp Ala Gly Ala Glu Leu
 690 695 700
 Tyr Leu Leu Thr Gly Glu Gln Ser Tyr Leu Gln Pro Leu Leu Ala Leu
 705 710 715 720
 Asp Thr Pro Ile Ser Ala Ala Ser Trp Ala Ser Val Ser Ala Leu Gly
 725 730 735
 Tyr Phe Ser Leu Ala Ser Ala Lys Gln Leu Glu Pro Ala Leu Arg Gln
 740 745 750
 Gln Val Gln Gln Lys Ile Gln Gln Ala Ala Ala Gln Ile Leu Gln Glu
 755 760 765
 His Gln Thr Ser Ala Tyr Gln Val Ala Met Thr Lys Asn Asp Phe Val
 770 775 780
 Trp Gly Ser Asn Ala Val Ala Met Asn Lys Ala Met Leu Leu Tyr Gln
 785 790 795 800
 Ala Trp Lys Ile Ala Pro Lys Pro Glu Leu Leu Gln Ala Met Gln Gly
 805 810 815
 Leu Val Asp Tyr Val Leu Gly Arg Asn Pro Leu Gln Gln Ser Tyr Val
 820 825 830
 Thr Gly Phe Gly Glu Gln Ser Pro Gln Gln Ile His His Arg Pro Ser
 835 840 845
 Ala Ala Asp Ala Ile Lys Ala Pro Val Pro Gly Trp Leu Val Gly Gly
 850 855 860
 Ala Gln Pro Gly Lys Gln Asp Lys Cys Thr Tyr Ala Gly Ala Leu Pro
 865 870 875 880
 Ala Val Gly Ala Leu Pro Ala Ala Ser Thr Leu Pro Ala Thr Thr Tyr
 885 890 895
 Leu Asp Asp Trp Cys Ser Tyr Ala Thr Asn Glu Val Ala Ile Asn Trp
 900 905 910
 Asn Ala Pro Leu Val Tyr Val Leu Ala Trp His Leu Ser Gln Asn Thr
 915 920 925
 Lys Thr Pro
 930

<210> 291

<211> 1230

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 291

atggtcaaag	aaagaagttt	tcttcatcat	tcattcaata	ggggcgaaaa	tggacaggac	60
agtctgatgt	ggaaaaaaga	ggcggatgat	cgaatctcag	agcatagaca	aagagatctt	120
gtgatcaacg	taacaaacgg	tgaaaaaaag	ccaatagcag	gtatagaggt	tgaaataaag	180
caaatcagac	atgaattcgc	ctttggttca	gcgatgaatg	atcaagtgtt	atttaatcaa	240
caatatgctg	attttttcgt	gaagtatttt	aattgggctg	tttttgaaaa	tgaggcaaaa	300
tggtatgcga	atgagccaca	aagagggaga	atcacctacg	aaaaagcaga	tgcgatgctg	360
aattttgcag	atcgacatca	gcttcacgtg	agagggcacg	ctttgttttg	ggaggtagag	420
gatgcgaatc	caagctggct	aagggtcactg	ccaaatcatg	aagtatatga	agccatgaaa	480
aaccggcttg	agcatgcggg	caatcacttt	aagggaaggt	tccgtcattg	ggatgtaaac	540

aatgaaatga	tgcatgggttc	atTTTTTaa	gatcgctttg	ggaaaaatat	ttggaagtgg	600
atgtatgaag	aaacgaaaaa	aattgaccct	caagcactat	tgtttgtgaa	tgattataat	660
gtgatctcat	atgggtgaaca	ccatgcctat	aaagcgcata	tcaatgaact	gcgtcagtta	720
ggcgcaccta	ttgaggcgat	tggggttcaa	ggccattttg	aagaacgggt	cgatccagtc	780
attgtcaaag	agagactcga	tgtgcttgct	gagctagggtc	ttccaatatg	ggtcacagag	840
tacgattcgg	ttcaccctga	ccctaatacga	agagcggata	acctggaagc	tttatatcgc	900
gtcgcattta	gtcatccagc	cgtaaaagga	gtgctgatgt	ggggattttg	ggcagggtgcc	960
cattggagag	gggaaaatgc	agccatcgtg	aattatgatt	ggctctttaa	tgaagcagga	1020
agacgttatg	aaaagcttct	aaatgagtgg	acgacccaaa	gaattgaaaa	aacagatgct	1080
aatggccatg	tgagatgtcc	agcatttcac	ggaacatatg	aggttcgaat	cggtaaagaa	1140
agtaaaatgt	tgaacagca	gacgattgaa	cttgattcaa	atgaacaaac	accgtttcaa	1200
ctagacgtga	tcctgcctca	agaaggatag				1230

<210> 292
 <211> 409
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 292

Met Val	Lys	Glu	Arg	Ser	Phe	Leu	His	His	Ser	Phe	Asn	Arg	Gly	Glu
1			5					10					15	
Asn	Gly	Gln	Asp	Ser	Leu	Met	Trp	Lys	Lys	Glu	Ala	Asp	Asp	Arg
			20					25				30		Ile
Ser	Glu	His	Arg	Gln	Arg	Asp	Leu	Val	Ile	Asn	Val	Thr	Asn	Gly
			35				40				45			Glu
Lys	Lys	Pro	Ile	Ala	Gly	Ile	Glu	Val	Glu	Ile	Lys	Gln	Ile	Arg
	50					55				60				His
Glu	Phe	Ala	Phe	Gly	Ser	Ala	Met	Asn	Asp	Gln	Val	Leu	Phe	Asn
65					70				75					80
Gln	Tyr	Ala	Asp	Phe	Phe	Val	Lys	Tyr	Phe	Asn	Trp	Ala	Val	Phe
			85						90					95
Asn	Glu	Ala	Lys	Trp	Tyr	Ala	Asn	Glu	Pro	Gln	Arg	Gly	Arg	Ile
			100					105					110	Thr
Tyr	Glu	Lys	Ala	Asp	Ala	Met	Leu	Asn	Phe	Ala	Asp	Arg	His	Gln
		115					120					125		Leu
Pro	Val	Arg	Gly	His	Ala	Leu	Phe	Trp	Glu	Val	Glu	Asp	Ala	Asn
	130					135					140			Pro
Ser	Trp	Leu	Arg	Ser	Leu	Pro	Asn	His	Glu	Val	Tyr	Glu	Ala	Met
145					150				155					160
Asn	Arg	Leu	Glu	His	Ala	Gly	Asn	His	Phe	Lys	Gly	Arg	Phe	Arg
			165						170					175
Trp	Asp	Val	Asn	Asn	Glu	Met	Met	His	Gly	Ser	Phe	Phe	Lys	Asp
			180					185					190	Arg
Phe	Gly	Lys	Asn	Ile	Trp	Lys	Trp	Met	Tyr	Glu	Glu	Thr	Lys	Lys
		195				200						205		Ile
Asp	Pro	Gln	Ala	Leu	Leu	Phe	Val	Asn	Asp	Tyr	Asn	Val	Ile	Ser
	210					215					220			Tyr
Gly	Glu	His	His	Ala	Tyr	Lys	Ala	His	Ile	Asn	Glu	Leu	Arg	Gln
225					230					235				240
Gly	Ala	Pro	Ile	Glu	Ala	Ile	Gly	Val	Gln	Gly	His	Phe	Glu	Glu
			245						250					255
Val	Asp	Pro	Val	Ile	Val	Lys	Glu	Arg	Leu	Asp	Val	Leu	Ala	Glu
		260						265					270	Leu
Gly	Leu	Pro	Ile	Trp	Val	Thr	Glu	Tyr	Asp	Ser	Val	His	Pro	Asp
		275					280					285		Pro
Asn	Arg	Arg	Ala	Asp	Asn	Leu	Glu	Ala	Leu	Tyr	Arg	Val	Ala	Phe
	290				295						300			Ser
His	Pro	Ala	Val	Lys	Gly	Val	Leu	Met	Trp	Gly	Phe	Trp	Ala	Gly
305					310					315				320
His	Trp	Arg	Gly	Glu	Asn	Ala	Ala	Ile	Val	Asn	Tyr	Asp	Trp	Ser
			325						330				335	Leu
Asn	Glu	Ala	Gly	Arg	Arg	Tyr	Glu	Lys	Leu	Leu	Asn	Glu	Trp	Thr
		340						345					350	Thr
Gln	Arg	Ile	Glu	Lys	Thr	Asp	Ala	Asn	Gly	His	Val	Arg	Cys	Pro
		355				360						365		Ala
Phe	His	Gly	Thr	Tyr	Glu	Val	Arg	Ile	Gly	Lys	Glu	Ser	Lys	Met
														Leu

370
 Lys Gln Gln Thr Ile Glu 375
 385 390 Leu Asp Ser Asn Glu 380
 Leu Asp Val Ile Leu Pro Gln Glu Gly 395
 400

<210> 293
 <211> 1002
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 293
 atgaagatga acagctccct cccctccctc cgcgatgtat tcgcgaatga tttccgcatc 60
 ggggcggcgg tcaatcctgt gacgatcgag atgcaaaaac agttgttgat cgatcatgtc 120
 aacagtatta cggcagagaa ccatatgaag tttgagcatc ttcagccgga agaaggga 180
 tttacctttc aggaagcgga tcggattgtg gatittgctt gttcgaccg aatggcggtt 240
 cgagggcaca cacttgatg gcacaaccag actccggatt ggggtgtttca agatgggtcaa 300
 ggccattttc tcagtcggga tgtgttgctt gagcggatga aatgtcacat ttcaactgtt 360
 gtacggcgat acaaggga aatatattgt tgggatgtca tcaacgaagc ggtagccgac 420
 gaaggagacg aattgttgag gccgtcgaag tggcgacaaa tcatcgggga cgattttatg 480
 gaacaagcat ttctctacgc ttatgaagct gaccagatg cactgctttt ttacaatgac 540
 tataatgaat gttttccgga aaagagagaa aaaatttttg cacttgctaa atcgctgcgt 600
 gataaaggca ttccgattca tggcatcggg atgcaagcgc attggagttt gactcgcccg 660
 tcgcttgatg aaattcgtgc ggccattgaa cgatatgcgt ccttggtgtg tgttctcat 720
 attacggaac tcgatgtatc catgtttgaa ttccacgac gtcgaaccga tttggcagct 780
 ccaacgtcag aaatgatcga acggcaggca gagcggatg ggcaaatttt tgctttgtt 840
 aaggagtatc gcgatgttat tcaaagtgtc acattttggg gaattgctga tgaccataca 900
 tggctcgata actttccagt gcacgggaga aaaaactggc cgcttttggt cgatgaacag 960
 cataaaccga aaccagcttt ttggcgggca gtgagtgtct ga 1002

<210> 294
 <211> 333
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 294
 Met Lys Met Asn Ser Ser Leu Pro Ser Leu Arg Asp Val Phe Ala Asn
 1 5 10 15
 Asp Phe Arg Ile Gly Ala Ala Val Asn Pro Val Thr Ile Glu Met Gln
 20 25 30
 Lys Gln Leu Leu Ile Asp His Val Asn Ser Ile Thr Ala Glu Asn His
 35 40 45
 Met Lys Phe Glu His Leu Gln Pro Glu Glu Gly Lys Phe Thr Phe Gln
 50 55 60
 Glu Ala Asp Arg Ile Val Asp Phe Ala Cys Ser His Arg Met Ala Val
 65 70 75 80
 Arg Gly His Thr Leu Val Trp His Asn Gln Thr Pro Asp Trp Val Phe
 85 90 95
 Gln Asp Gly Gln Gly His Phe Val Ser Arg Asp Val Leu Leu Glu Arg
 100 105 110
 Met Lys Cys His Ile Ser Thr Val Arg Arg Tyr Lys Gly Lys Ile
 115 120 125
 Tyr Cys Trp Asp Val Ile Asn Glu Ala Val Ala Asp Glu Gly Asp Glu
 130 135 140
 Leu Leu Arg Pro Ser Lys Trp Arg Gln Ile Ile Gly Asp Asp Phe Met
 145 150 155 160
 Glu Gln Ala Phe Leu Tyr Ala Tyr Glu Ala Asp Pro Asp Ala Leu Leu
 165 170 175
 Phe Tyr Asn Asp Tyr Asn Glu Cys Phe Pro Glu Lys Arg Glu Lys Ile
 180 185 190
 Phe Ala Leu Val Lys Ser Leu Arg Asp Lys Gly Ile Pro Ile His Gly
 195 200 205
 Ile Gly Met Gln Ala His Trp Ser Leu Thr Arg Pro Ser Leu Asp Glu

```

      210      215      220
Ile Arg Ala Ala Ile Glu Arg Tyr Ala Ser Leu Gly Val Val Leu His
225      230      235      240
Ile Thr Glu Leu Asp Val Ser Met Phe Glu Phe His Asp Arg Arg Thr
      245      250      255
Asp Leu Ala Ala Pro Thr Ser Glu Met Ile Glu Arg Gln Ala Glu Arg
      260      265      270
Tyr Gly Gln Ile Phe Ala Leu Phe Lys Glu Tyr Arg Asp Val Ile Gln
      275      280      285
Ser Val Thr Phe Trp Gly Ile Ala Asp Asp His Thr Trp Leu Asp Asn
      290      295      300
Phe Pro Val His Gly Arg Lys Asn Trp Pro Leu Leu Phe Asp Glu Gln
305      310      315      320
His Lys Pro Lys Pro Ala Phe Trp Arg Ala Val Ser Val
      325      330

```

<210> 295
 <211> 1134
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

```

<400> 295
atgagatccg tccgcatcgt cacctttgct ctcgccgccg cgctggccgt cccgctggtg      60
acgtcgacgg ccacggccaa gccgtccgcc gaccacgagg ccgcgccccca ctccaacgcc      120
aagttcgacc gcctgcgctg ggccgcccccc gaaggggttct tcataggctc cgcggcggcc      180
ggcgggcgcc accacctcga acaggactac ccggaccctt tcaccttcga caagaagtac      240
cggaagatcc tgggccagca gttcaactcg gtctccgccg agaaccagat gaagtgggag      300
ttcatccacc ccgagcgcgca ccagtaccgc ttcgaggagg ccgacgccat cgtcgagttc      360
gcccagcgga accgccaggc cgtgcgcggg cacaccctcc tgtggcacag ccagaacccc      420
gaatggctgg aggaggcgga cttcaccaag gaggaaactgc gcgccatcct caaggaccac      480
atcgacacgg tcgtcgggcg ctacgccggc aagatccagc agtgggacgt ggccaacgag      540
atcttcaacg accaggccga gctgcgcacc gacgagaaca tctggatacg tgagctcggc      600
ccggagatcg tcgcggacgc cttccgctgg gccacgaggg ccgaccccga ggccaagctg      660
ttctcaacg actacaacgt cgagggcacg aacgccaaga gcgacgccta ctacgagctc      720
gcccaggaga tgctggagca gggcggtgcc ctccacggat tcggcgccca gggccacctg      780
agcaccgcgt acggcttccc gggcgacctg cagcagaacc tgcagcgggt cgccgacctc      840
ggctctggaga ccgccatcac cgagatcgac gtccgcatgg acctcccggc gagcggcaag      900
cccaccaagg agcagctgcg gcagcaggcc gactactacc agcaggcact gtcggcctgc      960
ctggccgtga acgactgcaa ctcttcacc atctggggct tcaccgacaa gtactcgtgg      1020
gtgccggtct tcttcgaggg tgagggcagc gccacggta tgacggagaa gttcgtccgc      1080
aagccggcct tcttcgcctt gcagtcacc ctgaaggagg cgcgcaagcg ctga      1134

```

<210> 296
 <211> 377
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(26)

```

<400> 296
Met Arg Ser Val Arg Ile Val Thr Phe Ala Leu Ala Ala Ala Leu Ala
1      5      10      15
Val Pro Leu Val Thr Ser Thr Ala Thr Ala Lys Pro Ser Ala Asp His
      20      25      30
Glu Ala Ala Pro His Ser Asn Ala Lys Phe Asp Arg Leu Arg Trp Ala
      35      40      45
Ala Pro Glu Gly Phe Phe Ile Gly Ser Ala Ala Ala Gly Gly Gly His
      50      55      60
His Leu Glu Gln Asp Tyr Pro Asp Pro Phe Thr Phe Asp Lys Lys Tyr
65      70      75      80
Arg Lys Ile Leu Gly Gln Gln Phe Asn Ser Val Ser Ala Glu Asn Gln
      85      90      95

```

Met Lys Trp Glu Phe Ile His Pro Glu Arg Asp Gln Tyr Arg Phe Glu
 100 105 110
 Glu Ala Asp Ala Ile Val Glu Phe Ala Gln Arg Asn Arg Gln Ala Val
 115 120 125
 Arg Gly His Thr Leu Leu Trp His Ser Gln Asn Pro Glu Trp Leu Glu
 130 135 140
 Glu Gly Asp Phe Thr Lys Glu Glu Leu Arg Ala Ile Leu Lys Asp His
 145 150 155 160
 Ile Asp Thr Val Val Gly Arg Tyr Ala Gly Lys Ile Gln Gln Trp Asp
 165 170 175
 Val Ala Asn Glu Ile Phe Asn Asp Gln Ala Glu Leu Arg Thr Asp Glu
 180 185 190
 Asn Ile Trp Ile Arg Glu Leu Gly Pro Glu Ile Val Ala Asp Ala Phe
 195 200 205
 Arg Trp Ala His Glu Ala Asp Pro Glu Ala Lys Leu Phe Leu Asn Asp
 210 215 220
 Tyr Asn Val Glu Gly Ile Asn Ala Lys Ser Asp Ala Tyr Tyr Glu Leu
 225 230 235 240
 Ala Gln Glu Met Leu Glu Gln Gly Val Pro Leu His Gly Phe Gly Ala
 245 250 255
 Gln Gly His Leu Ser Thr Arg Tyr Gly Phe Pro Gly Asp Leu Gln Gln
 260 265 270
 Asn Leu Gln Arg Phe Ala Asp Leu Gly Leu Glu Thr Ala Ile Thr Glu
 275 280 285
 Ile Asp Val Arg Met Asp Leu Pro Ala Ser Gly Lys Pro Thr Lys Glu
 290 295 300
 Gln Leu Arg Gln Gln Ala Asp Tyr Tyr Gln Gln Ala Leu Ser Ala Cys
 305 310 315 320
 Leu Ala Val Asn Asp Cys Asn Ser Phe Thr Ile Trp Gly Phe Thr Asp
 325 330 335
 Lys Tyr Ser Trp Val Pro Val Phe Phe Glu Gly Glu Gly Ser Ala Thr
 340 345 350
 Val Met Thr Glu Lys Phe Val Arg Lys Pro Ala Phe Phe Ala Leu Gln
 355 360 365
 Ser Thr Leu Lys Glu Ala Arg Lys Arg
 370 375

<210> 297

<211> 1842

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 297

ttgaggctcag	gcgcggttctg	tttcatcata	gtcgtttttaa	tcctgaacct	tatatgcagg	60
gagttgtatg	agtgtaaaaa	agttgttacc	gcagcactgg	tatgcttggc	tttcgggtca	120
tcgctgactt	agggggcaatg	caccacattt	accaccagta	ccattcggaa	ttgcgatggg	180
atagattacg	agctctggag	ccagaataac	tctggcacga	ccaatatgca	aatacacggga	240
gggaactcga	atccaaacgg	tggaaacctt	gaggcgacat	ggagtggcac	gatcaatgtt	300
ctattccgcg	cgggtaaaaa	atggggcaca	tccagcacca	gtacccccaa	aaccatcggc	360
aatatctctc	ttgaattcgc	agcgacatgg	agttcggtcg	ataatgtgaa	aatgcttggc	420
atctatggct	gggctgatta	tccctcggga	agcgaaccaa	caaaaacgga	aagcgggtcaa	480
agcacgaact	tttccaatca	gattgagtat	tacatcattc	aagaccgcgg	tagctataac	540
ccggcatccg	gcggcaccaa	cgccaaaaag	tacggtgaag	ggacgatcga	cggaaatcgc	600
tatgattttt	atgtcgccga	ccgtatcggc	caggccatgc	tgacaggaac	gggaaatttc	660
aaacagtact	tcagcgtgcc	gaagagcaca	agcagtcaca	ggcaaagcgg	cacggtttcc	720
gtctccaaac	attttgaggc	ctgggaaaaa	gcgggcatga	agatgatgga	ttgtcgggta	780
tacgaagtgc	cgatgaaagt	ggaatcgtat	accggttccg	cgaatggcaa	cggctcggcg	840
aaagtgacca	aaaatctcct	cacgatcggc	ggaagcagca	gcaacgagtt	tagtctcgta	900
acgaatgttt	ctcctgtggt	cgcgggaaac	gtgtccaaga	gcccggacaa	cgcattcctat	960
gccccgaacg	cctccgttca	gctcacggcg	accccgaata	ccggttgga	gtttgtgggc	1020
tgggaagggg	acgcctcggg	ttccacgagc	ccaaccagcg	ttaccatgag	caaagacctc	1080
acggttacag	cgaagtttga	gctgggtatc	gaagaaggca	gcacaaacct	gatccaggat	1140
ggcaacttcc	cgaagcggcag	cgtaatctct	acagatgacg	gggcttcatg	gaaactcgga	1200
caaggggaaa	actggggaaa	ttccgcagcg	acaacgagcg	tcagcaatgg	aatcgcgaca	1260
gtcaatgtga	caactgtcgg	agcggaagct	tatcaaccgc	agcttgtaca	gtacgggttg	1320
ggactcgaca	tggacatgag	ttacaaactt	accttcaagg	caagagccga	tcgcgcaagg	1380

aagattgaag	ttgctgtcca	gcaggcgggtg	gatccttggg	ctgggttatgc	ttcccaggaa	1440
ttcgacctga	ccacgaccga	tcaggatttc	gagttcgtat	tcacgatgac	caacgccagc	1500
gacccggcat	cacagtccgc	gttcaatctt	ggccaggcga	caggcgatgt	ctatatcagt	1560
gatgttaaac	tggtatacac	gacaggcacc	acaccatat	cccgcacat	agtccgcggc	1620
aatacggcat	tcgtctcgg	aagtggcaga	accctgaata	tttcggcagt	cgacgcgtcc	1680
acacttcaga	tcaaggtagt	agatataaac	ggaaaggtaa	gagcgaattt	caacacggct	1740
ggtgcagcaa	gtgtttcctt	gtccaatatt	cctgcgggcc	agtacttcgt	tggtatcaca	1800
ggcacaggca	taaaacaaat	ctcaccgatc	gttttggaat	aa		1842

<210> 298

<211> 613

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 298

Met	Arg	Ser	Gly	Ala	Phe	Cys	Phe	Ile	Ile	Val	Val	Leu	Ile	Leu	Asn
1				5					10					15	
Leu	Ile	Cys	Arg	Glu	Leu	Tyr	Glu	Cys	Lys	Lys	Val	Val	Thr	Ala	Ala
			20					25					30		
Leu	Val	Cys	Leu	Ala	Phe	Gly	Ser	Ser	Leu	Thr	Trp	Gly	Gln	Cys	Thr
		35					40					45			
Thr	Phe	Thr	Thr	Ser	Thr	Ile	Arg	Asn	Cys	Asp	Gly	Ile	Asp	Tyr	Glu
	50					55					60				
Leu	Trp	Ser	Gln	Asn	Asn	Ser	Gly	Thr	Thr	Asn	Met	Gln	Ile	Thr	Gly
65				70					75					80	
Gly	Asn	Ser	Asn	Pro	Asn	Gly	Gly	Thr	Phe	Glu	Ala	Thr	Trp	Ser	Gly
			85						90					95	
Thr	Ile	Asn	Val	Leu	Phe	Arg	Ala	Gly	Lys	Lys	Trp	Gly	Thr	Ser	Ser
		100						105					110		
Thr	Ser	Thr	Pro	Lys	Thr	Ile	Gly	Asn	Ile	Ser	Leu	Glu	Phe	Ala	Ala
		115					120					125			
Thr	Trp	Ser	Ser	Val	Asp	Asn	Val	Lys	Met	Leu	Gly	Ile	Tyr	Gly	Trp
	130					135					140				
Ala	Tyr	Tyr	Pro	Ser	Gly	Ser	Glu	Pro	Thr	Lys	Thr	Glu	Ser	Gly	Gln
145					150				155					160	
Ser	Thr	Asn	Phe	Ser	Asn	Gln	Ile	Glu	Tyr	Tyr	Ile	Ile	Gln	Asp	Arg
			165					170						175	
Gly	Ser	Tyr	Asn	Pro	Ala	Ser	Gly	Gly	Thr	Asn	Ala	Lys	Lys	Tyr	Gly
		180						185					190		
Glu	Gly	Thr	Ile	Asp	Gly	Ile	Ala	Tyr	Asp	Phe	Tyr	Val	Ala	Asp	Arg
	195						200					205			
Ile	Gly	Gln	Ala	Met	Leu	Thr	Gly	Thr	Gly	Asn	Phe	Lys	Gln	Tyr	Phe
	210					215					220				
Ser	Val	Pro	Lys	Ser	Thr	Ser	Ser	His	Arg	Gln	Ser	Gly	Thr	Val	Ser
225					230				235					240	
Val	Ser	Lys	His	Phe	Glu	Ala	Trp	Glu	Lys	Ala	Gly	Met	Lys	Met	Met
			245						250					255	
Asp	Cys	Arg	Leu	Tyr	Glu	Val	Ala	Met	Lys	Val	Glu	Ser	Tyr	Thr	Gly
		260						265					270		
Ser	Ala	Asn	Gly	Asn	Gly	Ser	Ala	Lys	Val	Thr	Lys	Asn	Leu	Leu	Thr
		275					280					285			
Ile	Gly	Gly	Ser	Ser	Ser	Asn	Glu	Phe	Ser	Leu	Val	Thr	Asn	Val	Ser
	290					295					300				
Pro	Ala	Ser	Ala	Gly	Thr	Val	Ser	Lys	Ser	Pro	Asp	Asn	Ala	Ser	Tyr
305					310					315				320	
Ala	Pro	Asn	Ala	Ser	Val	Gln	Leu	Thr	Ala	Thr	Pro	Asn	Thr	Gly	Trp
			325						330					335	
Lys	Phe	Val	Gly	Trp	Glu	Gly	Asp	Ala	Ser	Gly	Ser	Thr	Ser	Pro	Thr
		340						345					350		
Ser	Val	Thr	Met	Ser	Lys	Asp	Leu	Thr	Val	Thr	Ala	Lys	Phe	Glu	Leu
		355					360					365			
Val	Ser	Glu	Glu	Gly	Ser	Thr	Asn	Leu	Ile	Gln	Asp	Gly	Asn	Phe	Pro
	370					375					380				
Ser	Gly	Ser	Val	Ile	Ser	Thr	Asp	Asp	Gly	Ala	Ser	Trp	Lys	Leu	Gly
385					390				395					400	
Gln	Gly	Glu	Asn	Trp	Gly	Asn	Ser	Ala	Ala	Thr	Thr	Ser	Val	Ser	Asn

Gly Ile Ala Thr Val Asn Val Thr Thr Val Gly Ala Glu Ala Tyr Gln
 405 420 425 430
 Pro Gln Leu Val Gln Tyr Gly Leu Gly Leu Asp Met Asp Met Ser Tyr
 435 440 445
 Lys Leu Thr Phe Lys Ala Arg Ala Asp Ala Ala Arg Lys Ile Glu Val
 450 455 460
 Ala Phe Gln Gln Ala Val Asp Pro Trp Ala Gly Tyr Ala Ser Gln Glu
 465 470 475 480
 Phe Asp Leu Thr Thr Thr Asp Gln Asp Phe Glu Phe Val Phe Thr Met
 485 490 495
 Thr Asn Ala Ser Asp Pro Ala Ser Gln Phe Ala Phe Asn Leu Gly Gln
 500 505 510
 Ala Thr Gly Asp Val Tyr Ile Ser Asp Val Lys Leu Val Tyr Thr Thr
 515 520 525
 Gly Thr Thr Pro Ile Ser Arg Thr Ile Val Arg Gly Asn Thr Ala Phe
 530 535 540
 Val Ser Val Ser Gly Arg Thr Leu Asn Ile Ser Ala Val Asp Ala Ser
 545 550 555 560
 Thr Leu Gln Ile Lys Val Val Asp Ile Asn Gly Lys Val Arg Ala Asn
 565 570 575
 Phe Asn Thr Ala Gly Ala Ala Ser Val Ser Leu Ser Asn Ile Pro Ala
 580 585 590
 Gly Gln Tyr Phe Val Gly Ile Thr Gly Thr Gly Ile Lys Gln Ile Ser
 595 600 605
 Pro Ile Val Leu Glu
 610

<210> 299

<211> 1047

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 299

atgtttttga	gtctcaaaag	agtggcggcg	cttgttttgcg	tgcgccgtct	cggcatctct	60
gcggcccaag	cacagacctg	cctgacctcg	agtcaaaccg	gcactaacia	cggcttctac	120
tattcgttct	ggaaagacaa	tcccggcacc	gtgaatttct	gtctgcagtc	cggcggccgc	180
tacacctcca	actggagcgg	catcaacaac	tgggtcggcg	gaaagggatg	gcagacgggt	240
tcccgcagag	tgggtgaacta	ctcgggcagc	ttcaattcgc	ctggcaatgg	gtacctgact	300
ctctatgggt	ggaccaccaa	tccgctcatc	gagtactaca	ttgtcgacaa	ctggggcagc	360
tatcgtccgc	cgggtgggca	gggtttcatg	ggcacggtga	ccagcgatgg	cgcgacgtat	420
gacgtctatc	gcacgcagcg	cgtaatcag	ccctgcatca	ccggcagcag	ttgcacgttc	480
tatcaatact	ggagcgtgcg	gcagtcgaag	cggaccggtg	gcacgatcac	caccggcaac	540
cacttcgatg	cctgggccag	ctacggaatg	aattctggcg	ctcacaacta	ccagatcatg	600
gcgaccgagg	gctatcaaag	cagcggcagc	tctgacatca	cggtgagtga	gggaagcagc	660
agcagtagca	gcgggtgggtg	cagcagcagc	agcagcagtg	gcggcggcgg	caccaagagc	720
ttcacgggtcc	gggcgcgcgg	aaccgcgggt	ggtaggtcca	tcacgctgcg	tgtgaacaat	780
cagaacgtgc	agacctggac	gctgggcacc	agcatgacga	actacacggc	atcgacgtcg	840
ttgagcgggtg	gcatcacctg	ggcttacacg	aacgacagtg	gcaatcgcg	cgtgcagggtg	900
gattacatca	tcgtgaacgg	ctcgacgcgt	cagtcagaag	cgcagagcta	caacaccggg	960
ctctatgcca	acggtagttg	tggtagcggc	tccaatagcg	aatggatgca	ttgcaacggc	1020
gccattggct	acgggaatac	gccgtag				1047

<210> 300

<211> 348

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(24)

<400> 300

Met Phe Leu Ser Leu Lys Arg Val Ala Ala Leu Val Cys Val Ala Gly

1 5 10 15
 Leu Gly Ile Ser Ala Ala Gln Ala Gln Thr Cys Leu Thr Ser Ser Gln
 20 25 30
 Thr Gly Thr Asn Asn Gly Phe Tyr Tyr Ser Phe Trp Lys Asp Asn Pro
 35 40 45
 Gly Thr Val Asn Phe Cys Leu Gln Ser Gly Gly Arg Tyr Thr Ser Asn
 50 55 60
 Trp Ser Gly Ile Asn Asn Trp Val Gly Gly Lys Gly Trp Gln Thr Gly
 65 70 75 80
 Ser Arg Arg Val Val Asn Tyr Ser Gly Ser Phe Asn Ser Pro Gly Asn
 85 90 95
 Gly Tyr Leu Thr Leu Tyr Gly Trp Thr Thr Asn Pro Leu Ile Glu Tyr
 100 105 110
 Tyr Ile Val Asp Asn Trp Gly Thr Tyr Arg Pro Pro Gly Gly Gln Gly
 115 120 125
 Phe Met Gly Thr Val Thr Ser Asp Gly Ala Thr Tyr Asp Val Tyr Arg
 130 135 140
 Thr Gln Arg Val Asn Gln Pro Cys Ile Thr Gly Ser Ser Cys Thr Phe
 145 150 155
 Tyr Gln Tyr Trp Ser Val Arg Gln Ser Lys Arg Thr Gly Gly Thr Ile
 165 170 175
 Thr Thr Gly Asn His Phe Asp Ala Trp Ala Ser Tyr Gly Met Asn Leu
 180 185 190
 Gly Ala His Asn Tyr Gln Ile Met Ala Thr Glu Gly Tyr Gln Ser Ser
 195 200 205
 Gly Ser Ser Asp Ile Thr Val Ser Glu Gly Ser Ser Ser Ser Ser
 210 215 220
 Gly Gly Gly Ser Ser Thr Ser Ser Ser Gly Gly Gly Gly Thr Lys Ser
 225 230 235
 Phe Thr Val Arg Ala Arg Gly Thr Ala Gly Gly Glu Ser Ile Thr Leu
 245 250 255
 Arg Val Asn Asn Gln Asn Val Gln Thr Trp Thr Leu Gly Thr Ser Met
 260 265 270
 Thr Asn Tyr Thr Ala Ser Thr Ser Leu Ser Gly Gly Ile Thr Val Ala
 275 280 285
 Tyr Thr Asn Asp Ser Gly Asn Arg Asp Val Gln Val Asp Tyr Ile Ile
 290 295 300
 Val Asn Gly Ser Thr Arg Gln Ser Glu Ala Gln Ser Tyr Asn Thr Gly
 305 310 315
 Leu Tyr Ala Asn Gly Ser Cys Gly Gly Gly Ser Asn Ser Glu Trp Met
 325 330 335
 His Cys Asn Gly Ala Ile Gly Tyr Gly Asn Thr Pro
 340 345

<210> 301

<211> 642

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 301

atgtttaagt	ttacaaagaa	attcttagtt	gggttaacgg	cagctttgat	gagtatgagc	60
ttgttttcgg	caaacgcctc	tcagactaac	acagactact	ggcaaaattg	gactgatggg	120
ggcggaacag	taaacgctgt	caatgggtct	ggcggaatt	acagtgtgaa	ttggtctaata	180
accggaatt	tcgttgttgg	ttaaagggttg	actacaggtt	cgccatttag	gacgataaac	240
tataatgccg	gagtttgggc	gccgaacggc	aatgcatatt	tgactttata	tggttgagcg	300
cgatccccctc	tcatagaata	ttatgtagtg	gattcatggg	gtacttatag	acctactgga	360
acgtataaag	gtacggttta	cagtgatggg	ggtacatatg	acgtgtacac	aactacacgt	420
tatgatgcac	cttccattga	tggcgataaa	actactttta	cgcagtactg	gagtggtcgc	480
cagtcgaaga	gaccaactgg	aagcaacgct	acaatcactt	tcagcaatca	cgtaaacgca	540
tggaagagat	atgggatgaa	tctgggtagt	aattgggtctt	accaagtctt	agcgacagag	600
ggatatcgaa	gtagtggaag	ttctaacgta	acagtgtggg	aa		642

<210> 302

<211> 213

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(28)

<400> 302

```

Met Phe Lys Phe Thr Lys Lys Phe Leu Val Gly Leu Thr Ala Ala Leu
 1      5      10      15
Met Ser Met Ser Phe Ser Ala Asn Ala Ser Ala Ala Asn Thr Asp
 20      25      30
Tyr Trp Gln Asn Trp Thr Asp Gly Gly Gly Thr Val Asn Ala Val Asn
 35      40      45
Gly Ser Gly Gly Asn Tyr Ser Val Asn Trp Ser Asn Thr Gly Asn Phe
 50      55      60
Val Val Gly Lys Gly Trp Thr Thr Gly Ser Pro Phe Arg Thr Ile Asn
 65      70      75      80
Tyr Asn Ala Gly Val Trp Ala Pro Asn Gly Asn Ala Tyr Leu Thr Leu
 85      90      95
Tyr Gly Trp Thr Arg Ser Pro Leu Ile Glu Tyr Tyr Val Val Asp Ser
100      105      110
Trp Gly Thr Tyr Arg Pro Thr Gly Thr Tyr Lys Gly Thr Val Tyr Ser
115      120      125
Asp Gly Gly Thr Tyr Asp Val Tyr Thr Thr Thr Arg Tyr Asp Ala Pro
130      135      140
Ser Ile Asp Gly Asp Lys Thr Thr Phe Thr Gln Tyr Trp Ser Val Arg
145      150      155      160
Gln Ser Lys Arg Pro Thr Gly Ser Asn Ala Thr Ile Thr Phe Ser Asn
165      170      175
His Val Asn Ala Trp Lys Arg Tyr Gly Met Asn Leu Gly Ser Asn Trp
180      185      190
Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Arg Ser Ser Gly Ser Ser
195      200      205
Asn Val Thr Val Trp
210

```

<210> 303

<211> 1404

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 303

```

ttgactataa aggcctggcc cgggacggct gccagctata ataccaatag aggttttatc      60
atgtcctacg ctacgttttaa gggggccgct accctagcga cgtccttcct gctcgcagtc      120
accttgacag cctgtggagg cagcaaatcc aaaccgcgtt tgccagaccc atcgaacagc      180
agctcgtcat caagcagcag ctcgtcatca agcagcagct cctcaagttc ctccagtagc      240
agttcagagct cttctagtgc tccctccagc caaacgttct tcattgagcc ggatttccag      300
cttcacaccc tggcggactt cccgattgga gtggcagctct cggcagccaa tgagccatac      360
agcatcttca accaaaccga tggtagtgat cggcaggatg tgatcctgga gcatttcaac      420
gaaatgaccg ctggcaacat catgaaaatg agctacgtgt acgcagggtca acgtgcaaat      480
cagcaacccg atcaattcga cttcagcaga gctgatgagc tggttgggtt tgcccacgca      540
aacagtgtga agattcacgg tcacgcccctc gtttggcacg ccgactatca agttccgggt      600
ttcatgcaga attatgatgg cgactttgct gagatgttgg ccaatcacgc gcggagtgtt      660
gtggaacatt ttgacgaaga gtttccaggt accgtgttca gctgggatgt ggtcaacgag      720
gcgataaccg acaacttcgg aaccgataca aatggctggc gccggtcgct gttttacaac      780
gcgctgccgc ccgcgcagaga agacgatatt cctgagtaca tccgcgttgc cttccaggcc      840
gctcgcgatg ccaaccggga catcgacctc tattacaatg attacgcaa taccgccaac      900
accaaccggc tgaacaaaac cctgcagatc gccgatgccc tggccgagga cgagctgatc      960
gacggtgtgg gattccagat gcacgtctat atgacgtacc cgagccttag tcacttccaa      1020
aacgcgtttc aagaagtggg tgatcgaggc ttgaagggtga agatcaccga gctggacgta      1080
tcggtgggtc acccatacgg tcagagcact ccgccaccgc agcccgctcta cgatgaagcg      1140
ttggcaggcg cacagaaaaa gcggttctgc gatataccca gacttatctt ggaaacgggt      1200
ccggctgagc ttcgcggcgg tctcactgtt tgggggcttg ccgacaacga aagctgggtt      1260
atgcaacagt tcaggaacgc aacgggagcg aactacaccg acgtgtggcc gttgtgtgtc      1320
aacgccgacc tgtcagccaa acctacactc caaggcgtgg ccgatgctct gcagggtctc      1380

```

ccctgcacca ccgacctcga ctaa

1404

<210> 304
 <211> 467
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(74)

<400> 304
 Met Thr Ile Lys Ala Trp Pro Gly Thr Ala Ala Ser Tyr Asn Thr Asn
 1 5 10 15
 Arg Gly Phe Ile Met Ser Tyr Ala Gln Phe Lys Gly Ala Ala Thr Leu
 20 25 30
 Ala Thr Ser Phe Leu Leu Ala Val Thr Leu Thr Ala Cys Gly Gly Ser
 35 40 45
 Lys Ser Lys Pro Val Leu Pro Asp Pro Ser Asn Ser Ser Ser Ser
 50 55 60
 Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser
 65 70 75 80
 Ser Ser Ser Ser Ser Ala Pro Ser Ser Gln Thr Phe Phe Ile Glu
 85 90 95
 Pro Asp Phe Gln Leu His Thr Leu Ala Asp Phe Pro Ile Gly Val Ala
 100 105 110
 Val Ser Ala Ala Asn Glu Pro Tyr Ser Ile Phe Asn Gln Thr Asp Gly
 115 120 125
 Thr Asp Arg Gln Asp Val Ile Leu Glu His Phe Asn Glu Met Thr Ala
 130 135 140
 Gly Asn Ile Met Lys Met Ser Tyr Val Tyr Ala Gly Gln Arg Ala Asn
 145 150 155 160
 Gln Gln Pro Asp Gln Phe Asp Phe Ser Arg Ala Asp Glu Leu Val Gly
 165 170 175
 Phe Ala His Ala Asn Ser Val Lys Ile His Gly His Ala Leu Val Trp
 180 185 190
 His Ala Asp Tyr Gln Val Pro Gly Phe Met Gln Asn Tyr Asp Gly Asp
 195 200 205
 Phe Ala Glu Met Leu Ala Asn His Ala Arg Ser Val Val Glu His Phe
 210 215 220
 Asp Glu Glu Phe Pro Gly Thr Val Val Ser Trp Asp Val Val Asn Glu
 225 230 235 240
 Ala Ile Thr Asp Asn Phe Gly Thr Asp Thr Asn Gly Trp Arg Arg Ser
 245 250 255
 Leu Phe Tyr Asn Ala Leu Pro Pro Ala Thr Glu Asp Asp Ile Pro Glu
 260 265 270
 Tyr Ile Arg Val Ala Phe Gln Ala Arg Asp Ala Asn Pro Asp Ile
 275 280 285
 Asp Leu Tyr Tyr Asn Asp Tyr Asp Asn Thr Ala Asn Thr Asn Arg Leu
 290 295 300
 Asn Lys Thr Leu Gln Ile Ala Asp Ala Leu Ala Glu Asp Glu Leu Ile
 305 310 315 320
 Asp Gly Val Gly Phe Gln Met His Val Tyr Met Thr Tyr Pro Ser Leu
 325 330 335
 Ser His Phe Gln Asn Ala Phe Gln Glu Val Val Asp Arg Gly Leu Lys
 340 345 350
 Val Lys Ile Thr Glu Leu Asp Val Ser Val Val Asn Pro Tyr Gly Gln
 355 360 365
 Ser Thr Pro Pro Pro Gln Pro Val Tyr Asp Glu Ala Leu Ala Gly Ala
 370 375 380
 Gln Lys Lys Arg Phe Cys Asp Ile Thr Arg Val Tyr Leu Glu Thr Val
 385 390 395 400
 Pro Ala Glu Leu Arg Gly Gly Leu Thr Val Trp Gly Leu Ala Asp Asn
 405 410 415
 Glu Ser Trp Leu Met Gln Gln Phe Arg Asn Ala Thr Gly Ala Asn Tyr
 420 425 430
 Thr Asp Val Trp Pro Leu Leu Phe Asn Ala Asp Leu Ser Ala Lys Pro

435
 Thr Leu Gln Gly Val Ala Asp Ala Leu Gln Gly Leu Pro Cys Thr Thr
 450
 Asp Leu Asp
 465

<210> 305
 <211> 3705
 <212> DNA
 <213> Bacteria

<400> 305
 atgaagagta ttgtaaacag agttgtatct atcgttacag ctttaataat gatttttggg 60
 acatcactgc tticacaaca cataagggca tttgctgatg acactaatac aaatctggtt 120
 tctaattggg accttgagac aggcacaatt gatggctgga ttaagcaagg taatcctaca 180
 ttagaagtaa cgactgaaca agcaattggg caatacagta tgaaagttac gggtagaaca 240
 cagacataaa aaggacctgc atatagcttt ttaggaaaaa tgcagaaagg tgaatcatat 300
 aatgtatcgc ttaaagttag acttgtttct ggacaaaatt cttctaattcc ttttattacc 360
 gtgactatgt tttagagaaga tgacaatggc aagcattatg atacaatagt ttggcaaaaa 420
 caagtttctg aagattcatg gactactgta agcgggactt atacattaga ttatactgga 480
 acattataaa cattatacat gtatgtagaa tcacccgata caacgctgga atactatatt 540
 gatgatgttg tagtgacacc acaaaatcca atacaagtag gaaatgtgat taccaatgga 600
 acttttgaat atggaaatac ttcaggatgg gttggaacag gctcatctgt tgtaaggca 660
 gtgtatggag tggctcatag cggaggttat agtttattga cgacagggag aacagctaatt 720
 tggaaatggc ctatgtatga tttgactggc aaaatagtag caggtaaca atacaatggt 780
 gatttttggg tgaaatttgt taatggcaat gatacagaac aaataaaggc tactgttaaa 840
 gcgacttcta acaaagacaa ttatatacaa gttaatgatt ttgtaaatgt aaataaaggc 900
 gaatggacag aaataaaggc cagttttact ttacctgtga cagattacag cgggtgtcagc 960
 atctatgtag aatctcaaaa tcctacttta gagttttaca ttgatgattt ttctgtaata 1020
 ggtgaaattt caaataatca gattacaata caaaatgata ttccggattt atattcagta 1080
 ttcaaagatt atttccccat cgggtgttgc gttgattcga gtagattaaa tgatgctgat 1140
 ccacatgctc aattgactgc taaacatttt aatatgcttg ttgcagaaaa tgccatgaaa 1200
 ccggaagctc tgcagcctac agaggggaaac tttacctttg ataatgctga taagattggt 1260
 gattatgaaa tagcacataa tatgaagatg agaggtcata cattgctttg gcataatcag 1320
 gttccggatt ggtttttcca ggacctatct gatccgtcta aaccagcttc aagggatctg 1380
 ctgcttcaaa gattaagaac gcacataaca actgtgttag atcattttaa aacaaaatac 1440
 ggttctcaaa atccaataat cggatgggat gttgtaaatg aggttcttga tgataatggc 1500
 aatttaagaa attctaagtg gttacaaatt ataggacctg attatataga aaaagctttt 1560
 gaatatgcgc atgaggcaga tccatctatg aaattgttta ttaatgatta caacatcgaa 1620
 aataatggcg ttaaaacaca ggcaatgtat gatttagtga aaaagttaaa aagtgaagg 1680
 gtgcctataa acggaatagg catgcaaatg cacataagca taaattcaaa tatagacaat 1740
 ataaaagctt ctatagaaaa acttgcatac ttaggtgtgg aaatacaggt aactgaatta 1800
 gatatgaaca tgaatggtaa tgtatctaac gacgcattgc ttaagcaagc gagattgtat 1860
 aaacaattat ttgacttatt taaagcagaa taactgctgt agtttttggg 1920
 ggagtttcag atgatgtaag ttggcttagt aagccaaatg ctccactact ttttgattca 1980
 aagttacagg caaagccagc atactgggca attgtagatc aaggcaaagc catacctgac 2040
 attcaatctg caaaagcttt agaaggatca ccgacgattg gtgcaaatgt tgatagttct 2100
 tggaaacttg taaaaccatt gtatgctaatt acttatgtga aaggaactat tggagcaact 2160
 gctgtctgtt ataataatct ggataactaaa aacttatatt tgttagtaca agtttcagac 2220
 aatactccat ctaataatga tggatcagag atttttgtgg ataagaatga caacaaatct 2280
 actacctatg aaagtgcaga tgaacattat atagtttaaga gggatggtag agggagtcca 2340
 aatattacaa agtatgtaat gtctaattgt gatggatatg tagcacagat agctattcca 2400
 attgaagaca ttatgctgtg gctgaatgat aaaattggat ttgatcag aataaatgat 2460
 gaccaaggca gtggcaatat aaatgcgata acagtttggg atgattatac aaacagtc 2520
 gatactaata cagcatattt tggagattta gtattatcaa aacctgcaca gattgcaaca 2580
 gctatatatg gcactcctgt tattgacggg aaagtagatg gtatttggaa taatgctgaa 2640
 gctatttcca caaatacatg ggtcttgggt tcatatattg gcagatgtaa tttaaataaa 2700
 atgtgggacg atccttatga acaggattct tgggaagttt ttgtagatca gaataatgat 2760
 tccagtgtga attatgaaaa tgatgatggg cagtttagag ttaactatga taatgaacaa 2820
 aagacaactt attatgaaaa ttcaaatgga tttaagtctg caacaagtct tacacaaaat 2880
 agttttggag gaagcactaa tttccttggg acgagtatta ctccgtcaaa tgggtactatc 2940
 ggaataattg tagaagaagc taacgatgca taacgaaatg gtaagaggac aggtattgtc 3000
 ataggttttg acttgcaggt aaattcttgg caagatactt ctggatttgg aaacttgatg 3060
 acatgggtgt atccaagcgg tgtccttaaa aaggttgggg catttaatga cataaaagac 3120
 cttacaggta agccatctgt tgccttaaaa aggttgggg acatagtaga agggatgaca 3180
 aattgggcga aagacgtaaa caagacagtg agcagagcag aatttacagc aatgatactg 3240
 gacacccagt atgaacaaaa agaagcatat agcggagaat ttagcgtatg aaaaagtggg 3300
 aggctattaa acataaaaga agaagcatat tacaacacag gaataatcga aggtgacgga 3360
 gactggtatg caaacgcgat agaagcagca tacaacacag gaataatcga aggtgacgga 3420

aagaacgcaa	ggccaaatga	cagcataaca	agagaagaga	tgacagcaat	agccatgagg	3480
gcatacgaga	tgctgacaca	gtacgaagaa	gagaatatag	gtgcgacaac	atttagcgac	3540
gcacaaatcca	taagcgattg	ggcaagaaat	gtagtggcaa	atgcagcgaa	attaggaata	3600
gtaaatggtg	agccaaataa	cgtatttgca	cctaaaggaa	atgccacaag	agcagaagca	3660
gcagctatca	tatacggctt	attagaaaaa	acaaataagc	tttaa		3705

<210> 306
 <211> 1234
 <212> PRT
 <213> Bacteria

<220>
 <221> SIGNAL
 <222> (1)...(32)

<400> 306

Met	Lys	Ser	Ile	Val	Asn	Arg	Val	Val	Ser	Ile	Val	Thr	Ala	Leu	Ile
1				5					10					15	
Met	Ile	Phe	Gly	Thr	Ser	Leu	Leu	Ser	Gln	His	Ile	Arg	Ala	Phe	Ala
			20					25					30		
Asp	Asp	Thr	Asn	Thr	Asn	Leu	Val	Ser	Asn	Gly	Asp	Phe	Glu	Thr	Gly
		35					40					45			
Thr	Ile	Asp	Gly	Trp	Ile	Lys	Gln	Gly	Asn	Pro	Thr	Leu	Glu	Val	Thr
	50				55						60				
Thr	Glu	Gln	Ala	Ile	Gly	Gln	Tyr	Ser	Met	Lys	Val	Thr	Gly	Arg	Thr
65					70				75					80	
Gln	Thr	Tyr	Glu	Gly	Pro	Ala	Tyr	Ser	Phe	Leu	Gly	Lys	Met	Gln	Lys
			85						90					95	
Gly	Glu	Ser	Tyr	Asn	Val	Ser	Leu	Lys	Val	Arg	Leu	Val	Ser	Gly	Gln
		100						105					110		
Asn	Ser	Ser	Asn	Pro	Phe	Ile	Thr	Val	Thr	Met	Phe	Arg	Glu	Asp	Asp
		115					120					125			
Asn	Gly	Lys	His	Tyr	Asp	Thr	Ile	Val	Trp	Gln	Lys	Gln	Val	Ser	Glu
	130					135					140				
Asp	Ser	Trp	Thr	Thr	Val	Ser	Gly	Thr	Tyr	Thr	Leu	Asp	Tyr	Thr	Gly
145					150					155					160
Thr	Leu	Lys	Thr	Leu	Tyr	Met	Tyr	Val	Glu	Ser	Pro	Asp	Pro	Thr	Leu
			165						170					175	
Glu	Tyr	Tyr	Ile	Asp	Asp	Val	Val	Val	Thr	Pro	Gln	Asn	Pro	Ile	Gln
		180						185					190		
Val	Gly	Asn	Val	Ile	Thr	Asn	Gly	Thr	Phe	Glu	Asn	Gly	Asn	Thr	Ser
		195					200					205			
Gly	Trp	Val	Gly	Thr	Gly	Ser	Ser	Val	Val	Lys	Ala	Val	Tyr	Gly	Val
	210					215					220				
Ala	His	Ser	Gly	Gly	Tyr	Ser	Leu	Leu	Thr	Thr	Gly	Arg	Thr	Ala	Asn
225					230					235					240
Trp	Asn	Gly	Pro	Ser	Tyr	Asp	Leu	Thr	Gly	Lys	Ile	Val	Pro	Gly	Gln
			245						250					255	
Gln	Tyr	Asn	Val	Asp	Phe	Trp	Val	Lys	Phe	Val	Asn	Gly	Asn	Asp	Thr
		260						265					270		
Glu	Gln	Ile	Lys	Ala	Thr	Val	Lys	Ala	Thr	Ser	Asn	Lys	Asp	Asn	Tyr
		275					280					285			
Ile	Gln	Val	Asn	Asp	Phe	Val	Asn	Val	Asn	Lys	Gly	Glu	Trp	Thr	Glu
	290					295					300				
Ile	Lys	Gly	Ser	Phe	Thr	Leu	Pro	Val	Thr	Asp	Tyr	Ser	Gly	Val	Ser
305					310					315					320
Ile	Tyr	Val	Glu	Ser	Gln	Asn	Pro	Thr	Leu	Glu	Phe	Tyr	Ile	Asp	Asp
			325						330					335	
Phe	Ser	Val	Ile	Gly	Glu	Ile	Ser	Asn	Asn	Gln	Ile	Thr	Ile	Gln	Asn
		340						345					350		
Asp	Ile	Pro	Asp	Leu	Tyr	Ser	Val	Phe	Lys	Asp	Tyr	Phe	Pro	Ile	Gly
	355						360					365			
Val	Ala	Val	Asp	Ser	Ser	Arg	Leu	Asn	Asp	Ala	Asp	Pro	His	Ala	Gln
		370				375					380				
Leu	Thr	Ala	Lys	His	Phe	Asn	Met	Leu	Val	Ala	Glu	Asn	Ala	Met	Lys
385					390					395					400
Pro	Glu	Ser	Leu	Gln	Pro	Thr	Glu	Gly	Asn	Phe	Thr	Phe	Asp	Asn	Ala
			405						410					415	
Asp	Lys	Ile	Val	Asp	Tyr	Glu	Ile	Ala	His	Asn	Met	Lys	Met	Arg	Gly

420 425 430
 His Thr Leu Trp His Asn Gln Val Pro Asp Trp Phe Phe Gln Asp
 435 440 445
 Pro Ser Asp Pro Ser Lys Pro Ala Ser Arg Asp Leu Leu Leu Gln Arg
 450 455 460
 Leu Arg Thr His Ile Thr Thr Val Leu Asp His Phe Lys Thr Lys Tyr
 465 470 475 480
 Gly Ser Gln Asn Pro Ile Ile Gly Trp Asp Val Val Asn Glu Val Leu
 485 490 495
 Asp Asp Asn Gly Asn Leu Arg Asn Ser Lys Trp Leu Gln Ile Ile Gly
 500 505 510
 Pro Asp Tyr Ile Glu Lys Ala Phe Glu Tyr Ala His Glu Ala Asp Pro
 515 520 525
 Ser Met Lys Leu Phe Ile Asn Asp Tyr Asn Ile Glu Asn Asn Gly Val
 530 535 540
 Lys Thr Gln Ala Met Tyr Asp Leu Val Lys Lys Leu Lys Ser Glu Gly
 545 550 555 560
 Val Pro Ile Asn Gly Ile Gly Met Gln Met His Ile Ser Ile Asn Ser
 565 570 575
 Asn Ile Asp Asn Ile Lys Ala Ser Ile Glu Lys Leu Ala Ser Leu Gly
 580 585 590
 Val Glu Ile Gln Val Thr Glu Leu Asp Met Asn Met Asn Gly Asn Val
 595 600 605
 Ser Asn Asp Ala Leu Leu Lys Gln Ala Arg Leu Tyr Lys Gln Leu Phe
 610 615 620
 Asp Leu Phe Lys Ala Glu Lys Gln Tyr Ile Thr Ala Val Val Phe Trp
 625 630 635 640
 Gly Val Ser Asp Asp Val Ser Trp Leu Ser Lys Pro Asn Ala Pro Leu
 645 650 655
 Leu Phe Asp Ser Lys Leu Gln Ala Lys Pro Ala Tyr Trp Ala Ile Val
 660 665 670
 Asp Gln Gly Lys Ala Ile Pro Asp Ile Gln Ser Ala Lys Ala Leu Glu
 675 680 685
 Gly Ser Pro Thr Ile Gly Ala Asn Val Asp Ser Ser Trp Lys Leu Val
 690 695 700
 Lys Pro Leu Tyr Ala Asn Thr Tyr Val Lys Gly Thr Ile Gly Ala Thr
 705 710 715 720
 Ala Ala Val Lys Ser Met Trp Asp Thr Lys Asn Leu Tyr Leu Leu Val
 725 730 735
 Gln Val Ser Asp Asn Thr Pro Ser Asn Asn Asp Gly Ile Glu Ile Phe
 740 745 750
 Val Asp Lys Asn Asp Asn Lys Ser Thr Thr Tyr Glu Ser Asp Asp Glu
 755 760 765
 His Tyr Ile Val Lys Arg Asp Gly Thr Gly Ser Ser Asn Ile Thr Lys
 770 775 780
 Tyr Val Met Ser Asn Ala Asp Gly Tyr Val Ala Gln Ile Ala Ile Pro
 785 790 795 800
 Ile Glu Asp Ile Ser Pro Val Leu Asn Asp Lys Ile Gly Phe Asp Ile
 805 810 815
 Arg Ile Asn Asp Asp Gln Gly Ser Gly Asn Ile Asn Ala Ile Thr Val
 820 825 830
 Trp Asn Asp Tyr Thr Asn Ser Gln Asp Thr Asn Thr Ala Tyr Phe Gly
 835 840 845
 Asp Leu Val Leu Ser Lys Pro Ala Gln Ile Ala Thr Ala Ile Tyr Gly
 850 855 860
 Thr Pro Val Ile Asp Gly Lys Val Asp Gly Ile Trp Asn Asn Ala Glu
 865 870 875 880
 Ala Ile Ser Thr Asn Thr Trp Val Leu Gly Ser Asn Gly Ala Thr Ala
 885 890 895
 Thr Ala Lys Met Met Trp Asp Asp Lys Tyr Leu Tyr Ile Leu Ala Asp
 900 905 910
 Val Thr Asp Asn Asn Leu Asn Lys Ser Ser Val Asn Pro Tyr Glu Gln
 915 920 925
 Asp Ser Val Glu Val Phe Val Asp Gln Asn Asn Asp Lys Thr Thr Tyr
 930 935 940
 Tyr Glu Asn Asp Asp Gly Gln Phe Arg Val Asn Tyr Asp Asn Glu Gln
 945 950 955 960
 Ser Phe Gly Gly Ser Thr Asn Ser Asn Gly Phe Lys Ser Ala Thr Ser
 965 970 975

Leu Thr Gln Asn Gly Tyr Ile Val Glu Glu Ala Ile Pro Trp Thr Ser
 980 985 990
 Ile Thr Pro Ser Asn Gly Thr Ile Ile Gly Phe Asp Leu Gln Val Asn
 995 1000 1005
 Asp Ala Asp Glu Asn Gly Lys Arg Thr Gly Ile Val Thr Trp Cys Asp
 1010 1015 1020
 Pro Ser Gly Asn Ser Trp Gln Asp Thr Ser Gly Phe Gly Asn Leu Met
 1025 1030 1035 1040
 Leu Thr Gly Lys Pro Ser Gly Val Leu Lys Lys Val Val Ala Phe Asn
 1045 1050 1055
 Asp Ile Lys Asp Asn Trp Ala Lys Asp Val Ile Glu Val Leu Ala Ser
 1060 1065 1070
 Arg His Ile Val Glu Gly Met Thr Asp Thr Gln Tyr Glu Pro Asn Lys
 1075 1080 1085
 Thr Val Thr Arg Ala Glu Phe Thr Ala Met Ile Leu Arg Leu Leu Asn
 1090 1095 1100
 Ile Lys Glu Glu Ala Tyr Ser Gly Glu Phe Ser Asp Val Lys Ser Gly
 1105 1110 1115 1120
 Asp Trp Tyr Ala Asn Ala Ile Glu Ala Ala Tyr Lys Thr Gly Ile Ile
 1125 1130 1135
 Glu Gly Asp Gly Lys Asn Ala Arg Pro Asn Asp Ser Ile Thr Arg Glu
 1140 1145 1150
 Glu Met Thr Ala Ile Ala Met Arg Ala Tyr Glu Met Leu Thr Gln Tyr
 1155 1160 1165
 Glu Glu Glu Asn Ile Gly Ala Thr Thr Phe Ser Asp Asp Lys Ser Ile
 1170 1175 1180
 Ser Asp Trp Ala Arg Asn Val Val Ala Asn Ala Ala Lys Leu Gly Ile
 1185 1190 1195 1200
 Val Asn Gly Glu Pro Asn Asn Val Phe Ala Pro Lys Gly Asn Ala Thr
 1205 1210 1215
 Arg Ala Glu Ala Ala Ala Ile Ile Tyr Gly Leu Leu Glu Lys Thr Asn
 1220 1225 1230
 Lys Leu

<210> 307
 <211> 3729
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 307
 atgaaatccg ccacgaggcga cgctgcgttt atggcggcga tggaacgcaa ctccgacatc 60
 gtgacgatgc agtgctacgc acctatcttt gtcaatgtca atcccggcgg gcggcagtgg 120
 cgccccaat tgatgggcta cgatgcgtta agcgctttcg gctcggccctc gtattacgcc 180
 atcaaaatgt tcagcaacaa tttgggcgat acgattttga agcccagtc cagcgggtgcg 240
 cgctgccag tttccgttac acaagagcag aaaagcggca cgattttcat taaattgggtg 300
 aaccgcgaaa cgacgccaca gagcgtaaaa attgatctca aaggcgtgcg ctccgtcgaa 360
 ttcagcggca ccgccactgt ttagctgcc gactccggcg cgcttaactc cattgatgcg 420
 cctaccaaag tcgttcctgt cacgcgcaga attactggaa tcagccccctc gtttgcgcaa 480
 acgctggagc cgtattcgat tactgttttg caaatcaagg ccactgctct gccaacggcg 540
 acagcaaacg ccgttgccgc gccaaccttc accacggagc caaaagtga caccaccag 600
 cccgttacta tttccgttgc gacttcgcag ccgcagccag tttcgcaacc gtcgccgat 660
 gcaaacgcta tcgcgcgcgt gaaaacgct ttcaaaggca agttcctcat tggcaccgtg 720
 ttgagcggcg cagacctgcg cggccagcaa acgcgcagcg tgggtatcgc caccacgat 780
 ttcgacgcct ttacggcaga aaacgaaatg aagccggatg cgatgcagcc tcgcgaagg 840
 caattcaact ttgtgcgcg cgaccgctta gtggaactgg ccgaaaaaag cggcgccacg 900
 cccatcggcc acacgctaatt ctggcactcg caaacaccac gctgggtctt tgaagggccg 960
 gacggacaac cggcgaatcg tgaattggcc ttggcgcgga tgcgcaagca tattgcaacc 1020
 gtcgtcggcc actacaaagg gcgcgtcaag cagtgggacg tgggtgaacga agccataaac 1080
 gatggccccg gcgtgtctgc ccaaagcccg tggctgcgcg ctatcggcga agactacatc 1140
 gctgaagcgt ttcgagccgc gcacgcgcgc gaccctgacg caattctcat ctataacgat 1200
 tacaacatcg aaatgggcta caaacggccc aaagcaatcc aactcttaaa atctctggtt 1260
 gatcagaag tgccgattca tgccgttggc attcagggcc actggcgtat ggacaccaac 1320
 ctcaccgaag tggacaagc tattaaagaa ttctcgccgc tgggcctgaa ggtgatgatc 1380
 actgaactcg acatcggtgt tttgccgacg cgttatcagg gcgctgatat ttcgcagggtg 1440
 caaaacatga cgcctgaaca gcgcgcgcgc gtgaaccat ataccaatgg tttgcccgcac 1500

```

gacgtagcgc aaaaacacgc cgataaatat cgccaggcct tcgatatttt cctgcgctac 1560
aaagatgtca tcgaacgcgt cactgttctgg ggtgtggacg atgctcattc gtggctgaac 1620
ggtttcccga ttccgggcg caccgattac ccattgcttt tcgaccggca gggcaaacct 1680
aaacccgcct ttttcgccgt gcaaaacctg gctttaggcg tgaccgccgc gccgcaatcg 1740
aatgtttcgt ctgcacccag agctgtcgct caagctgcgc cggcaaccag caatattcgt 1800
ggccaggaat ttccacgggt cgaaagcgat ttgctgtgta cgtttcgcat caaagcccc 1860
gaagcgcaaa aagtgcagtt cgatttaggc aagccttacg acgcgacccg cgatgctgaa 1920
ggcaactgga cggcaaccac cgagccgcag gtgcccggct ttcactatta caatttggtg 1980
attgacggcg tgcgcgtgaa cgacccggcc agcgaaacct tttacggcg cggccggccag 2040
atgagcggca tcgaaatccc cgaccctgac agcgcttttt attcgccgca aaacgtgccc 2100
cacggcgaag tgcgcgagcg ctggtatttt tccaacacca cgaggcggtg gcgtcgcat 2160
ttcattttata cgccgcccgg ttatgacacg aatcaagtgg agcgttttcc cgttttgat 2220
ttacagcatg gcggcggcga agatgagcgc ggctggcctc aacagggtcg catgagcttt 2280
atcatggata atctcatcgc cacacgtaaa gccaaaccga tgctgtgggt gatggaacaa 2340
ggctatgcgc gcaagcccaa cgagccgcag gtgcccgttc gcccgcccgg cggtagcgcc 2400
ggagccatgc ccccgatttt caatcgcatg ttcggcacac tgggtgaagt gttcaccaa 2460
gacctgattc cctttattga tgctaattac cgcaaaaaa ccgaccgcga aaaccgcgcc 2520
atggccggac ttctgatggg cgggatgcaa agtttcctta ttggctgtc gaacaccgat 2580
ttattcgcgc acatcgaggg cttagcggc gcgggcccgg gtttcgggtg cggcaccttc 2640
gacgcacaaa cggcgacgg cgggtgtaat gccgacgccg acgcgttcaa caaaaaagtt 2700
cgacgctgt ttctcagcat cggcacagcg gaaaacgaac gctttcagag cagcgtgcgc 2760
ggttaccgcg atgcgtgac caaagcgggc atcaaaaacca cattctacga atcgccggc 2820
acctcacatg aatggctgac atggcggcgt agcctgaaag aattcgacc gctctgttt 2880
caagaagtcg aagtgcaaat tgagcgcggc ccgaatgcgc gcccaattgc gccgcaaccg 2940
attaatctcg gccccgacga taaacccgca ttcccccg tggccggcgg tttcgatgtc 3000
cgccgcaacg atattccgca tggcgaaatc aaactcgtgg aatatccatc cgctacagt 3060
ggcacaatc gtaagatgca ggtttacacg ccgcccgggt acaatccgca agaaaagtat 3120
gcgggtgctt atctgttgca cggaaatcggc ggcgacgagt gggaatggaa aaatggcggc 3180
acgcggaag tgattctcga taatctttac gccgcgaaaa aactccagcc gatgatttg 3240
gtcatgccca atggccgcgc gcaaaaagat gaccgtccaa tcggcaatgt gttcgcatc 3300
gcgcctgctt ttgaaacctt cgagaaagat ttactcaacg acgtaattcc gtttatcgaa 3360
aagaattatc cagtcaaaac cggcgcggaa aatcgcgcgc ttgcgggcct ttcgatggg 3420
ggtgggcaat cgctgaactt tggctctggg aatctcgaca cttttgcgtg ggtcggcggc 3480
ttttcttcgg cgccaacac gcgcaactgc tagccaatcc cgacgatgcc 3540
aaaaagaagc tgaaattggt atgggtttcg tgtggcgata aagacggctt gttttcatc 3600
agccagcgca cgcacgcta tctcgccgaa aacaatgtgc cacacgtctg gcatgtgcag 3660
cccggcggcc acgatttccg agtgtggaag caagaccttt ataacttttc gcaactgctg 3720
ttccgctaa 3729

```

<210> 308

<211> 1242

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 308

```

Met Lys Ser Ala Ile Gly Asp Ala Ala Phe Met Ala Ala Met Glu Arg
1      5      10      15
Asn Ser Asp Ile Val Thr Met Gln Cys Tyr Ala Pro Ile Phe Val Asn
20     25     30
Val Asn Pro Gly Gly Arg Gln Trp Arg Pro Asn Leu Met Gly Tyr Asp
35     40     45
Ala Leu Ser Ala Phe Gly Ser Pro Ser Tyr Tyr Ala Ile Lys Met Phe
50     55     60
Ser Asn Asn Leu Gly Asp Thr Ile Leu Lys Pro Ser Leu Ser Gly Ala
65     70     75     80
Arg Leu Pro Val Ser Val Thr Gln Glu Gln Lys Ser Gly Thr Ile Phe
85     90     95
Ile Lys Leu Val Asn Pro Gln Thr Thr Pro Gln Ser Val Lys Ile Asp
100    105    110
Leu Lys Gly Val Arg Ser Val Glu Phe Ser Gly Thr Ala Thr Val Leu
115    120    125
Ala Ala Asp Ser Gly Ala Leu Asn Ser Ile Asp Ala Pro Thr Lys Val
130    135    140
Val Pro Val Thr Arg Arg Ile Thr Gly Ile Ser Pro Ser Phe Ala Gln
145    150    155    160
Thr Leu Glu Pro Tyr Ser Ile Thr Val Leu Gln Ile Lys Ala Thr Ala

```

Leu	Pro	Thr	Ala	165	Thr	Ala	Asn	Ala	Val	170	Ala	Pro	Pro	Thr	Phe	175	Thr	Thr
Glu	Pro	Lys	Val	180	Asn	Thr	Thr	Thr	Pro	Val	185	Thr	Ile	Pro	Val	190	Ala	Thr
Ser	Gln	Pro	Gln	195	Pro	Val	Ser	Gln	Pro	Ser	Pro	Asp	205	Ala	Asn	Ala	Ile	
Ala	Pro	Leu	Lys	210	Asn	Ala	Phe	Lys	Gly	Lys	Phe	Leu	Ile	Gly	Thr	Val	240	
Leu	Ser	Gly	Pro	225	Asp	Leu	Arg	Gly	Gln	Gln	Thr	Arg	Ser	Val	Gly	Ile	255	
Ala	Thr	Thr	His	245	Asp	Ala	Phe	Thr	250	Ala	Glu	Asn	Glu	Met	Lys	Pro	270	
Asp	Ala	Met	Gln	260	Pro	Arg	Glu	Gly	265	Gln	Phe	Asn	Phe	Ala	Ala	Gly	Asp	
Arg	Leu	Val	Glu	275	Leu	Ala	Glu	Lys	280	Ser	Gly	Ala	Thr	Pro	Ile	Gly	His	
Thr	Leu	Ile	Trp	290	His	Ser	Gln	Thr	295	Pro	Arg	Trp	Phe	Phe	Glu	Gly	Pro	
Asp	Gly	Gln	Pro	305	Asn	Arg	Glu	Leu	310	Ala	Leu	Ala	Arg	Met	Arg	Lys	320	
His	Ile	Ala	Thr	325	Val	Val	Gly	His	330	Tyr	Lys	Gly	Arg	Val	Lys	Gln	Trp	
Asp	Val	Val	Asn	340	Glu	Ala	Ile	Asn	345	Asp	Gly	Pro	Gly	Val	Leu	Arg	Gln	
Ser	Pro	Trp	Leu	355	Arg	Ala	Ile	Gly	360	Glu	Asp	Tyr	Ile	Ala	Glu	Ala	Phe	
Arg	Ala	Ala	His	370	Ala	Ala	Asp	Pro	375	Asp	Ala	Ile	Leu	Ile	Tyr	Asn	Asp	
Tyr	Asn	Ile	Glu	385	Met	Gly	Tyr	Lys	390	Arg	Pro	Lys	Ala	Ile	Gln	Leu	Leu	
Lys	Ser	Leu	Val	405	Asp	Gln	Lys	Val	410	Pro	Ile	His	Ala	Val	Gly	Ile	Gln	
Gly	His	Trp	Arg	420	Met	Asp	Thr	Asn	425	Leu	Thr	Glu	Val	Glu	Gln	Ala	Ile	
Lys	Glu	Phe	Ser	435	Ala	Leu	Gly	Leu	440	Lys	Val	Met	Ile	Thr	Glu	Leu	Asp	
Ile	Gly	Val	Leu	450	Pro	Thr	Arg	Tyr	455	Gln	Gly	Ala	Asp	Ile	Ser	Gln	Val	
Gln	Asn	Met	Thr	465	Pro	Glu	Gln	Arg	470	Ala	Ala	Val	Asn	Pro	Tyr	Thr	Asn	
Gly	Leu	Pro	Asp	485	Asp	Val	Ala	Gln	490	Lys	His	Ala	Asp	Lys	Tyr	Arg	Gln	
Ala	Phe	Asp	Ile	500	Phe	Leu	Arg	Tyr	505	Lys	Asp	Val	Ile	Glu	Arg	Val	Thr	
Phe	Trp	Gly	Val	515	Asp	Asp	Ala	His	520	Ser	Trp	Leu	Asn	Gly	Phe	Pro	Ile	
Arg	Gly	Arg	Thr	530	Asp	Tyr	Pro	Leu	535	Leu	Phe	Asp	Arg	Gln	Gly	Lys	Pro	
Lys	Pro	Ala	Phe	545	Phe	Ala	Val	Gln	550	Asn	Leu	Ala	Leu	Gly	Val	Thr	Ala	
Ala	Pro	Gln	Ser	565	Asn	Ala	Ser	Ser	570	Ala	Pro	Arg	Ala	Val	Ala	Gln	Ala	
Ala	Pro	Ala	Thr	580	Ser	Asn	Ile	Arg	585	Gly	Gln	Glu	Phe	Pro	Arg	Val	Glu	
Ser	Asp	Leu	Arg	595	Val	Thr	Phe	Arg	600	Ile	Lys	Ala	Pro	Glu	Ala	Gln	Lys	
Val	Gln	Phe	Asp	610	Leu	Gly	Lys	Pro	615	Tyr	Asp	Ala	Thr	Arg	Asp	Ala	Glu	
Gly	Asn	Trp	Thr	625	Ala	Thr	Thr	Glu	630	Pro	Gln	Val	Pro	Gly	Phe	His	Tyr	
Tyr	Asn	Leu	Val	645	Ile	Asp	Gly	Val	650	Arg	Val	Asn	Asp	Pro	Ala	Ser	Glu	
Thr	Phe	Tyr	Gly	660	Ala	Gly	Arg	Gln	665	Met	Ser	Gly	Ile	Glu	Ile	Pro	Asp	
Pro	Asp	Ser	Ala	675	Phe	Tyr	Ser	Pro	680	Gln	Asn	Val	Pro	His	Gly	Glu	Val	
Arg	Glu	Arg	Trp	690	Tyr	Phe	Ser	Asn	695	Thr	Thr	Gln	Ala	Trp	Arg	Arg	Ile	
				705					710									720

Phe Ile Tyr Thr Pro Gly Tyr Asp Thr Asn Gln Val Glu Arg Phe
 725 730 735
 Pro Val Leu Tyr Leu Gln His Gly Gly Gly Glu Asp Glu Arg Gly Trp
 740 745 750
 Pro Gln Gln Gly Arg Met Ser Phe Ile Met Asp Asn Leu Ile Ala Thr
 755 760 765
 Arg Lys Ala Lys Pro Met Leu Val Val Met Glu Gln Gly Tyr Ala Arg
 770 775 780
 Lys Pro Asn Glu Pro Gln Val Pro Leu Arg Pro Pro Gly Gly Ser Ala
 785 790 795 800
 Gly Ala Met Pro Pro Asp Phe Asn Arg Met Phe Gly Thr Leu Gly Glu
 805 810 815
 Val Phe Thr Lys Asp Leu Ile Pro Phe Ile Asp Ala Asn Tyr Arg Thr
 820 825 830
 Lys Thr Asp Arg Glu Asn Arg Ala Met Ala Gly Leu Ser Met Gly Gly
 835 840 845
 Met Gln Ser Phe Leu Ile Gly Leu Ser Asn Thr Asp Leu Phe Ala His
 850 855 860
 Ile Gly Gly Phe Ser Gly Ala Gly Gly Gly Phe Gly Gly Gly Thr Phe
 865 870 875 880
 Asp Ala Lys Thr Ala His Gly Gly Val Met Ala Asp Ala Asp Ala Phe
 885 890 895
 Asn Lys Lys Val Arg Thr Leu Phe Leu Ser Ile Gly Thr Ala Glu Asn
 900 905 910
 Glu Arg Phe Gln Ser Ser Val Arg Gly Tyr Arg Asp Ala Leu Thr Lys
 915 920 925
 Ala Gly Ile Lys Thr Thr Phe Tyr Glu Ser Pro Gly Thr Ser His Glu
 930 935 940
 Trp Leu Thr Trp Arg Arg Ser Leu Lys Glu Phe Ala Pro Leu Leu Phe
 945 950 955 960
 Gln Glu Val Glu Val Gln Ile Glu Arg Gly Pro Asn Ala Arg Pro Ile
 965 970 975
 Ala Pro Gln Pro Ile Asn Leu Gly Pro Asp Asp Lys Pro Ala Phe Pro
 980 985 990
 Pro Val Pro Ala Gly Phe Asp Val Arg Arg Asn Asp Ile Pro His Gly
 995 1000 1005
 Glu Ile Lys Leu Val Glu Tyr Pro Ser Ala Thr Val Gly Thr Asn Arg
 1010 1015 1020
 Lys Met Gln Val Tyr Thr Pro Pro Gly Tyr Asn Pro Gln Glu Lys Tyr
 1025 1030 1035 1040
 Ala Val Leu Tyr Leu Leu His Gly Ile Gly Gly Asp Glu Trp Glu Trp
 1045 1050 1055
 Lys Asn Gly Gly Thr Pro Glu Val Ile Leu Asp Asn Leu Tyr Ala Ala
 1060 1065 1070
 Lys Lys Leu Gln Pro Met Ile Val Val Met Pro Asn Gly Arg Ala Gln
 1075 1080 1085
 Lys Asp Asp Arg Pro Ile Gly Asn Val Phe Ala Ser Ala Pro Ala Phe
 1090 1095 1100
 Glu Thr Phe Glu Lys Asp Leu Leu Asn Asp Val Ile Pro Phe Ile Glu
 1105 1110 1115 1120
 Lys Asn Tyr Pro Val Lys Thr Gly Ala Glu Asn Arg Ala Leu Ala Gly
 1125 1130 1135
 Leu Ser Met Gly Gly Gly Gln Ser Leu Asn Phe Gly Leu Gly Asn Leu
 1140 1145 1150
 Asp Thr Phe Ala Trp Val Gly Gly Phe Ser Ser Ala Pro Asn Thr Arg
 1155 1160 1165
 Thr Gly Ala Arg Leu Leu Ala Asn Pro Asp Asp Ala Lys Lys Lys Leu
 1170 1175 1180
 Lys Leu Leu Trp Val Ser Cys Gly Asp Lys Asp Gly Leu Phe Phe Ile
 1185 1190 1195 1200
 Ser Gln Arg Thr His Arg Tyr Leu Ala Glu Asn Asn Val Pro His Val
 1205 1210 1215
 Trp His Val Gln Pro Gly Gly His Asp Phe Arg Val Trp Lys Gln Asp
 1220 1225 1230
 Leu Tyr Asn Phe Ser Gln Leu Leu Phe Arg
 1235 1240

<210> 309

<211> 1830

<212> DNA
<213> Unknown

<220>
<223> obtained from an environmental sample.

<400> 309

ttgaaaaaac	tcacaatcgc	cctatccctt	gcaatcactt	ttgccgcgcc	agtttttgcg	60
acagatgctt	gtttgcaaaa	tactcaatta	aatgctaccg	cccaaggagc	acaaacctgg	120
actggcaaaa	aaggagctac	aacttttaggc	ggttcagggtg	acgatgctta	tggagttgaa	180
acttggacag	aagctgggtg	agacgtact	aaatttacat	ggtttgacc	aaatcagggt	240
ggtggtttcg	cttatagagc	ggaatggaca	aattccacag	attacttagg	tcgctttggt	300
tatttttggg	gtattgacgg	gaaaaaatgg	gacaaattag	gagacctttg	cgttgattat	360
aactataaaa	gatctgccaa	tggcactgga	ggttcatatt	cttatatagg	cgtttatgga	420
tggacaaacg	ctggcggtgg	tactgaagct	gaatattata	tagttgaaga	ttggtttgga	480
gaaaatcaac	agactgcaaa	taatttaggc	aatggttgcc	aagagcacgg	tgaaattaca	540
gtggacgaga	aaagctataa	agttgtcact	tgcataagac	cagcaggctc	tggctgcgta	600
acttgaacg	gacaacaatt	cgggcaagtc	tttagcatac	gccaaaggcat	gagaagcgaa	660
caacctaaaa	catgcggaac	aatctccatc	aaaaagcact	ttgaagaatg	ggtaaaaatg	720
acaaccgaaa	aaagcggaca	atccccagct	aaatatattt	acgataaaac	ttatgaatcc	780
aaatttttag	cggaggcaca	aggcggcact	ggttggcttg	aaaccacttt	ccttaaattt	840
tctagaaacg	gtgattgcgg	cttcgatatt	cctgatggtc	atttcaccct	tcaattagct	900
acctctccgt	ctgaggcgcg	tactgtaagc	agaaacccaa	aggaatcttc	ttatgcctcc	960
ggttcaactg	taactctcac	agccactccg	gcagcagggt	ggaaatttgc	cagttggagt	1020
ggtgatgcat	tgcaaacacc	aagcccattg	acagtcacta	tggacaaaaa	caaggtaatc	1080
acagcgaaat	ttacccccgt	tgtagatctc	aataaaaaac	ttgttacaaa	tggtactttc	1140
acgaataaag	agagttggac	ttttaatact	ggttccagtt	atggcaactc	cgaaggaact	1200
tttgatgtgt	caaacagcga	aggcagaata	aatgtgacga	aaattggctc	caaccctggg	1260
gaaccacagc	tcgttcaaaa	tggcatcacc	cttgttgaag	gaatgaatta	caaaaataact	1320
tttgaagctt	ctgcctctgc	tgctcgaaaa	ataggccttg	ttatacaaat	ggcaggcggc	1380
gattatacca	cttattttga	aaaagatata	gacttgactg	cctcaaatca	acaattttcg	1440
tatgaattca	aaatgaatgc	accaagcgat	gaaagcggtc	gtattgggtt	taatcttggg	1500
caaacgactg	gaaacgtcac	tctaagcaaa	atcacccttc	attattttaga	agaagattca	1560
cacgaaccgt	ccgacaatcc	ctgcgaagac	ccaagtccaa	ttttgaaaaa	acgcattccct	1620
gcaactcatt	cttcctttca	aacgcttagc	gacaaagcct	tgcgcataga	agtgaacgct	1680
ccaaccattg	tggacatttt	tgacctgaga	gggaataaag	ttaagagttt	gaatgtctcc	1740
ggctcgcaaa	cggttaaatt	atccctgcc	agcggagtgt	attttgccaa	agcacgtgga	1800
atgaaaagcg	tcagatttgt	gttgaggtaa				1830

<210> 310
<211> 609
<212> PRT
<213> Unknown

<220>
<223> obtained from an environmental sample.

<221> SIGNAL
<222> (1)...(20)

<400> 310

Met	Lys	Lys	Leu	Thr	Ile	Ala	Leu	Ser	Leu	Ala	Ile	Thr	Phe	Ala	Ala
1				5				10					15		
Pro	Val	Phe	Ala	Thr	Asp	Ala	Cys	Leu	Gln	Asn	Thr	Gln	Leu	Asn	Ala
			20				25					30			
Thr	Ala	Gln	Gly	Ala	Gln	Thr	Trp	Thr	Gly	Lys	Lys	Gly	Ala	Thr	Thr
		35				40					45				
Leu	Gly	Gly	Ser	Gly	Asp	Asp	Ala	Tyr	Gly	Val	Glu	Thr	Trp	Thr	Glu
	50				55				60						
Ala	Gly	Gly	Asp	Ala	Thr	Lys	Phe	Thr	Trp	Phe	Gly	Pro	Asn	Gln	Gly
65				70				75					80		
Gly	Gly	Phe	Ala	Tyr	Arg	Ala	Glu	Trp	Thr	Asn	Ser	Thr	Asp	Tyr	Leu
			85					90					95		
Gly	Arg	Phe	Gly	Tyr	Phe	Trp	Gly	Ile	Asp	Gly	Lys	Lys	Trp	Asp	Lys
		100					105					110			
Leu	Gly	Asp	Leu	Cys	Val	Asp	Tyr	Asn	Tyr	Lys	Arg	Ser	Ala	Asn	Gly
		115				120					125				
Thr	Gly	Gly	Ser	Tyr	Ser	Tyr	Ile	Gly	Val	Tyr	Gly	Trp	Thr	Asn	Ala
	130					135					140				

<400> 311

atgCGGaaaa	gagtaataagc	tttattttgta	actctcatct	ttgtcatgtc	tattttaagt	60
ccaggatatac	ttccattttct	gagtactaaa	gcaaatgtct	aaacacaaaa	tacaccaaca	120
atttttaaaat	ttgatttttga	aagcggtaat	caaggctgga	cggggagagg	tctttcaaca	180
actgttgcaa	ccgtttacaa	tgttgcttat	gaagggtgatt	attcattaaa	agtttctggc	240
agaaatgctt	catgggatgg	agctgttatt	gatttaacag	acaagctttc	ggcaaatgtg	300
agttatacag	tttctctgtt	tgttcgtcac	agtgaccaaa	aacctcaaag	attttcagta	360
tatgcatatg	taaaagattc	agcaagtga	aaatatattc	cagttgtaga	taaagttgca	420
gttcctaatt	attggaagca	actggttggt	aaattcacaa	tcaacacttc	aaatccagtc	480
caaaagattc	agctgcttgt	ctgtgttcct	acaaataaat	cattggaatt	ttttatcgat	540
agcgtattaa	ttgcaagtag	tgCaggagca	acatctggag	ttgtaaaatc	cacaaatttt	600
gaaagcggta	caacagaagg	ctggcaagca	aggggaacag	gttctgttgc	tcagattagt	660
gttggtttcta	cagtagctca	ttcaggtagt	aaaagtgtgt	atgtgacagg	gcgagttcaa	720
acgtggcaag	gtgcacaaat	agatttgaca	agtttggttag	agaagggtaa	agaatatcag	780
ttttctgtgt	gggtatatca	ggatagtggg	agcgaccaaa	agctgacact	gaccatggaa	840
aggaaaaatg	cagatggaaag	tacaaattat	gatacaataa	aatggcagca	aacagtttca	900
agcaatacat	gggtagagct	aacaggttca	tatacagtac	ctgcaacagc	aacacaacta	960
atattctaca	ttgaatcacc	caatgctacc	ctaagctttt	atattgatga	ttttactgct	1020
gttgataaaa	atgcaccagt	tgtagcgcct	ggaattataa	aatcagccac	atttgaaagc	1080
ggtcacacag	aagactggca	agcaagaggg	acaggagtga	ccgtttctgt	tgtaaacaca	1140
gtggcacata	ctgggagcaa	gagtttgtat	gtgacaggga	gaagtcaaaa	ttggcatggg	1200
gcagaaattg	atctgacaaa	tgTgctagag	aagggaagg	aatatcaatt	ttctgtgtgg	1260
gtatatcagg	atagtggaaag	cgatcagaag	ctgacattga	ccatgcaaaag	gaaaaatgca	1320
gataacacaa	cagattatga	ttctataaaa	tatcagcaga	cagttagcaac	aaatacatgg	1380
gtagagctaa	caggttcata	tacagtgccg	acaacagcca	cccagttaat	tttatatggt	1440
gaagctgcag	atactaccct	aagcttttat	attgatgatt	ttactgctgt	tgataaaaaac	1500
ccagaggtaa	taccaacagt	atcgagagta	ccagaatggg	aaattccttc	actctttgag	1560
cagtatacga	attattttcag	cattgggtgtg	gcaataccat	ataaagtact	tacaaatcca	1620
accgaaaagg	caatgggtact	caaacatttc	aacagtataa	cagctgagaa	tgaaatgaaa	1680
ctgtatgtca	ttcaaaagac	agaagggaat	tttacattta	atgttgCaga	ccaatacgta	1740
gattttgcac	agcaaaatag	aattggaatc	agagggtcaca	ctcttgtttg	gcatcagcaa	1800
acaccaaatt	ggttctttcca	gcatagtgtat	ggtactccgc	ttgatccaag	caatcctgct	1860
gacaaacaac	ttctacgcga	tagattaaga	acgcatatcc	aaacacttgt	tggaagatac	1920
gcagggaaaa	tttagtcata	ggatgttgta	aacgagggca	ttgacgagaa	ccagccagat	1980
ggatacagaa	gaagtgaatg	gtacagaata	ttggggccaa	ctgatacaac	agatggcatt	2040
ccagaatata	ttctgcttgc	attccagtat	gcaagagagg	cagacccaaa	tactaagctc	2100
ttttataacg	actataatac	agaaaaatcca	aagaaaagac	agtttatata	caatcttggt	2160
aagaagctca	aagaaagagg	cttgattgat	ggtgtaggct	tgCagtgcc	tattaatgtc	2220
gattcaccta	cagttaaaga	gatagaggat	acaattaaac	tgTtagtac	aatccctggc	2280
ttagacattc	acattacaga	gcttgacatt	agcgtttata	caagcagcag	ccagagatat	2340
gatactcttc	cacaggatat	aatgataaaa	caggctttga	agttcaaaga	gctttttgag	2400
atgctaaaaa	gatatagcta	tgTgtcaca	aacgttactt	tctggggact	caaagatgac	2460
tattctggc	tttcaacaag	cagatctaac	tgggccactac	tgTttgcaa	caactaccag	2520
gcaaaatttg	catactgggc	aattgttgaa	ccgtcagtat	tgccacttgc	tataaacaac	2580
ggatatgcaa	acaatgcata	agcaaggata	gatggagttt	tagacagaga	atacaaaggt	2640
gcgattccaa	ttagatttac	aaatgaaagt	ggacaagaag	ttgcaactgt	tcgagctcta	2700
tggaaattcaa	gtgaactcag	cctctatata	tcgggtcaatg	atacaacaat	agatgctgct	2760
aatgataaag	tagttgtatt	tgtagaccag	gataatggaa	aaatgccaga	aattaaacct	2820
gatgactatt	gggtttcaat	tacgagaact	ggtaacaaag	cacaatcagc	tcaaggctat	2880
gtaaaggatt	atgctgtcgt	gcagcaagca	aatggatatg	tggttgagtt	gaagctttta	2940
attaataaca	cgttactgt	taactcttct	ataggttttg	atatagcaat	ctttgacaat	3000
ggagttcaat	acagtggaa	tgacaagaca	aactcacagt	ttatagaaac	tgataactat	3060
ggtatttttaa	caatggcaga	tagcgtcaag	tttgcttctg	ctccaaaagg	tacagcaata	3120
attgatgcag	aattagatga	tacatggaaa	aacgctcagg	aaataacaac	tgacacaaag	3180
gtcacggtta	caggcacagt	atacgactca	gcttatgcaa	aggctaagat	gatgtgggat	3240
gaaaatagta	tctatgtcta	tgcaattgtt	tatgacttgc	ttttgaacaa	ggctaataca	3300
aatccatggg	agcagatttc	aattgagata	tttggtgtag	aaaataatca	caaaacgcct	3360
tactatgaaa	atgatgatgt	tcagtacaga	gtgaactatg	agaatactca	aacatttggc	3420
acgaacggtg	ctcctcagaa	cttcattaca	gcaacaaaga	taattccaaa	cggatatata	3480
gtggaagctc	aagttttacat	gaggacgaca	aagctttctg	aagggaatggt	tataggcttt	3540
gacattcaag	tgaattgatgc	agaccataca	ggtaaaagag	tcggtgttct	aacctggaat	3600
gataagggtg	ggaacaatta	tagagacaca	acaagggttta	gatgcttaga	gcttgtagca	3660
gcacctgtaa	gccagccacc	aatacaagct	ccatcaccat	cacaaccaac	aacaataacg	3720
tatatactaa	caccgacacc	aacacagcca	tcaaccctaa	cacagcagca	acctgctcag	3780
caaccatcac	agcagcaaca	gcaaccgcaa	cagcagcagc	ctgcacagac	acaacaacct	3840
cagacacagc	ctgcacaaaa	gcctcagaat	gttgtttcga	taaagataga	ccagacaaaa	3900
gctgagaacat	ttactgttgg	cgctgatacc	aagggtgttg	tacctcaagg	ttctgttaact	3960
ggtgcaaaact	ga					3972

<210> 312
 <211> 1323
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(33)

<400> 312
 Met Arg Lys Arg Val Ile Ala Leu Phe Val Thr Leu Ile Phe Val Met
 1 5 10 15
 Ser Ile Leu Ser Pro Gly Tyr Leu Pro Phe Leu Ser Thr Lys Ala Asn
 20 25 30
 Ala Gln Thr Gln Asn Thr Pro Thr Ile Leu Lys Phe Asp Phe Glu Ser
 35 40 45
 Gly Asn Gln Gly Trp Thr Gly Arg Gly Leu Ser Thr Thr Val Ala Thr
 50 55 60
 Val Tyr Asn Val Ala Tyr Glu Gly Asp Tyr Ser Leu Lys Val Ser Gly
 65 70 75 80
 Arg Asn Ala Ser Trp Asp Gly Ala Val Ile Asp Leu Thr Asp Lys Leu
 85 90 95
 Ser Ala Asn Val Ser Tyr Thr Val Ser Leu Phe Val Arg His Ser Asp
 100 105 110
 Gln Lys Pro Gln Arg Phe Ser Val Tyr Ala Tyr Val Lys Asp Ser Ala
 115 120 125
 Ser Glu Lys Tyr Ile Pro Val Val Asp Lys Val Ala Val Pro Asn Tyr
 130 135 140
 Trp Lys Gln Leu Val Gly Lys Phe Thr Ile Asn Thr Ser Asn Pro Val
 145 150 155 160
 Gln Lys Ile Gln Leu Val Cys Val Pro Thr Asn Lys Ser Leu Glu
 165 170 175
 Phe Phe Ile Asp Ser Val Leu Ile Ala Ser Ser Ala Gly Ala Thr Ser
 180 185 190
 Gly Val Val Lys Ser Thr Asn Phe Glu Ser Gly Thr Thr Glu Gly Trp
 195 200 205
 Gln Ala Arg Gly Thr Gly Ser Val Ala Gln Ile Ser Val Val Ser Thr
 210 215 220
 Val Ala His Ser Gly Ser Lys Ser Leu Tyr Val Thr Gly Arg Val Gln
 225 230 235 240
 Thr Trp Gln Gly Ala Gln Ile Asp Leu Thr Ser Leu Leu Glu Lys Gly
 245 250 255
 Lys Glu Tyr Gln Phe Ser Val Trp Val Tyr Gln Asp Ser Gly Ser Asp
 260 265 270
 Gln Lys Leu Thr Leu Thr Met Glu Arg Lys Asn Ala Asp Gly Ser Thr
 275 280 285
 Asn Tyr Asp Thr Ile Lys Trp Gln Gln Thr Val Ser Ser Asn Thr Trp
 290 295 300
 Val Glu Leu Thr Gly Ser Tyr Thr Val Pro Ala Thr Ala Thr Gln Leu
 305 310 315 320
 Ile Phe Tyr Ile Glu Ser Pro Asn Ala Thr Leu Ser Phe Tyr Ile Asp
 325 330 335
 Asp Phe Thr Ala Val Asp Lys Asn Ala Pro Val Val Ala Pro Gly Ile
 340 345 350
 Ile Lys Ser Ala Thr Phe Glu Ser Gly Thr Thr Glu Asp Trp Gln Ala
 355 360 365
 Arg Gly Thr Gly Val Thr Val Ser Val Val Asn Thr Val Ala His Thr
 370 375 380
 Gly Ser Lys Ser Leu Tyr Val Thr Gly Arg Ser Gln Asn Trp His Gly
 385 390 395 400
 Ala Glu Ile Asp Leu Thr Asn Val Leu Glu Lys Gly Lys Glu Tyr Gln
 405 410 415
 Phe Ser Val Trp Val Tyr Gln Asp Ser Gly Ser Asp Gln Lys Leu Thr
 420 425 430
 Leu Thr Met Gln Arg Lys Asn Ala Asp Asn Thr Thr Asp Tyr Asp Ser
 435 440 445
 Ile Lys Tyr Gln Gln Thr Val Ala Thr Asn Thr Trp Val Glu Leu Thr
 Page 234

Page 235

Lys Thr Asn Ser Gln Phe Ile Glu Thr Asp Asn Tyr Gly Ile Leu Thr
 1010 1015 1020
 Met Ala Asp Ser Val Lys Phe Ala Ser Ala Pro Lys Gly Thr Ala Ile
 1025 1030 1035 1040
 Ile Asp Ala Glu Leu Asp Asp Thr Trp Lys Asn Ala Gln Glu Ile Thr
 1045 1050 1055
 Thr Asp Thr Lys Val Thr Val Thr Gly Thr Val Tyr Asp Ser Ala Tyr
 1060 1065 1070
 Ala Lys Ala Lys Met Met Trp Asp Glu Asn Ser Ile Tyr Val Tyr Ala
 1075 1080 1085
 Ile Val Tyr Asp Leu Leu Leu Asn Lys Ala Asn Thr Asn Pro Trp Glu
 1090 1095 1100
 Gln Asp Ser Ile Glu Ile Phe Val Asp Glu Asn Asn His Lys Thr Pro
 1105 1110 1115 1120
 Tyr Tyr Glu Asn Asp Val Gln Tyr Arg Val Asn Tyr Glu Asn Thr
 1125 1130 1135
 Gln Thr Phe Gly Thr Asn Gly Ala Pro Gln Asn Phe Ile Thr Ala Thr
 1140 1145 1150
 Lys Ile Ile Pro Asn Gly Tyr Ile Val Glu Ala Gln Val Tyr Met Arg
 1155 1160 1165
 Thr Thr Lys Leu Ser Glu Gly Met Val Ile Gly Phe Asp Ile Gln Val
 1170 1175 1180
 Asn Asp Ala Asp His Thr Gly Lys Arg Val Gly Val Leu Thr Trp Asn
 1185 1190 1195 1200
 Asp Lys Val Gly Asn Asn Tyr Arg Asp Thr Thr Arg Phe Arg Cys Leu
 1205 1210 1215
 Glu Leu Val Ala Ala Pro Val Ser Gln Pro Pro Ile Gln Ala Pro Ser
 1220 1225 1230
 Pro Ser Gln Pro Thr Thr Ile Thr Tyr Ile Leu Thr Pro Thr Pro Thr
 1235 1240 1245
 Gln Pro Ser Thr Gln Thr Gln Gln Gln Pro Ala Gln Gln Pro Ser Gln
 1250 1255 1260
 Gln Gln Gln Gln Pro Gln Gln Gln Gln Pro Ala Gln Thr Gln Gln Pro
 1265 1270 1275 1280
 Gln Thr Gln Pro Ala Gln Lys Pro Gln Asn Val Val Ser Ile Lys Ile
 1285 1290 1295
 Asp Gln Thr Lys Ala Glu Thr Phe Thr Val Gly Ala Asp Thr Lys Val
 1300 1305 1310
 Val Val Pro Gln Gly Ser Val Thr Gly Ala Asn
 1315 1320

<210> 313
 <211> 1392
 <212> DNA
 <213> Bacteria

<400> 313
 gtgaaacgtc tatccgcgct gaccgcccgtc gtattgttag cgctaacaac tcacgtcggc 60
 gccgctgacc ccgcgccacc cgccaccggc cccgccatcg acttccgggc cgaactccag 120
 cccatcgacg gattcggcctt ctccatggcc ttccagcggg ccgacctgct gcacggcgcg 180
 cgcgccctca gccccgccaa gcggcgcgag gtgctcgacc tgctgctcga caaggagagg 240
 ggcgcggggc tgcgatcct gcgctgggc atcgggtcgt cgaccgaccg ggtctacgac 300
 cacatgccga cgatcctgcc gaccgatccc ggcgggcccg acgccccgcc gaagtacgtc 360
 tgggacggct gggacggcgg ccaggtctgg ctcgccaagg aggccaaagg gtacggcgctc 420
 aagcggttct tcgcccacgc ctggagcgcg ccggccttca tgaagaccaa cggcagcgag 480
 aacgacggcg gcgagctccg gccggaatgg cgccaggcct acgcgaacta cctcgtcaag 540
 tacgcgaagt tctaccaacg ggaaggcatc ccgatcaccg acctgggggt caccacgaa 600
 cccgactggg cggcgaccta gcctcgatg cgtttacc cgcagcaggc cgtcgacttc 660
 ctcaagggtc tcgggcccac cgtccgcgcg tccggactga agaccggcgt cgtctgctgc 720
 gacgcggcgg gctgggaccg gcaggtcgcc tacaccgagg ccatcgaggc ggaccccgag 780
 gccgccaaag ccgtgcggag cgtcaccggc caccgtaca gcggtccgac cacggtcccg 840
 cagcccaccg acaagcgggt ctggatgtcg gagtggtcac cggacggcac cacttggaac 900
 gagaactggg acgacggcag cggctacgac ggcctcaccg tcgccgccga catccagaac 960
 accctcaccg tcggcaacgc caacgcctac gtctactgga ccggcgcgct cctcggcgcc 1020
 accgggggac tcatccagct cgccaacccc ggcgactcct accgggtgtc caagcggtac 1080
 tgggcgctgg ccgccttcag ccgcttcac cgccttcacg accggtcacg accggtcacg 1140
 aacgccgacc cggccctgag cgtcacggcc ttccgcaaca ccgacggcag ccgctgtatc 1200
 gagatcctca acacggcgac caccgagaag tccgccaggt tcgccctccg cggcggccac 1260
 gaccggcacc ccgaggggcta cgtcaccgac gagaccgct cgatcacc cggccacgct 1320

gcctccgcgc gcggtacgac cctcaaggcc acgctcgccc cgcgcgcgct gaccacgac 1380
gtcctcgact ga 1392

<210> 314
<211> 463
<212> PRT
<213> Bacteria

<220>
<221> SIGNAL
<222> (1)...(22)

<400> 314

Met	Lys	Arg	Leu	Ser	Ala	Leu	Thr	Ala	Val	Val	Leu	Leu	Ala	Leu	Thr
1				5					10					15	
Thr	His	Val	Ala	Ala	Ala	Asp	Pro	Ala	Pro	Pro	Ala	Thr	Gly	Pro	Ala
			20					25					30		
Ile	Asp	Phe	Arg	Ala	Glu	Leu	Gln	Pro	Ile	Asp	Gly	Phe	Gly	Phe	Ser
		35					40					45			
Met	Ala	Phe	Gln	Arg	Ala	Asp	Leu	Leu	His	Gly	Ala	Arg	Gly	Leu	Ser
	50					55					60				
Pro	Ala	Lys	Arg	Arg	Glu	Val	Leu	Asp	Leu	Leu	Leu	Asp	Lys	Glu	Arg
65					70				75					80	
Gly	Ala	Gly	Leu	Ser	Ile	Leu	Arg	Leu	Gly	Ile	Gly	Ser	Ser	Thr	Asp
				85					90					95	
Arg	Val	Tyr	Asp	His	Met	Pro	Thr	Ile	Leu	Pro	Thr	Asp	Pro	Gly	Gly
			100					105					110		
Pro	Asp	Ala	Pro	Pro	Lys	Tyr	Val	Trp	Asp	Gly	Trp	Asp	Gly	Gly	Gln
		115					120					125			
Val	Trp	Leu	Ala	Lys	Glu	Ala	Lys	Ala	Tyr	Gly	Val	Lys	Arg	Phe	Phe
		130				135					140				
Ala	Asp	Ala	Trp	Ser	Ala	Pro	Ala	Phe	Met	Lys	Thr	Asn	Gly	Ser	Glu
145					150					155					160
Asn	Asp	Gly	Gly	Glu	Leu	Arg	Pro	Glu	Trp	Arg	Gln	Ala	Tyr	Ala	Asn
				165					170					175	
Tyr	Leu	Val	Lys	Tyr	Ala	Lys	Phe	Tyr	Gln	Arg	Glu	Gly	Ile	Pro	Ile
			180					185					190		
Thr	Asp	Leu	Gly	Phe	Thr	Asn	Glu	Pro	Asp	Trp	Ala	Ala	Thr	Tyr	Ala
		195					200					205			
Ser	Met	Arg	Phe	Thr	Pro	Gln	Gln	Ala	Val	Asp	Phe	Leu	Lys	Val	Leu
	210					215					220				
Gly	Pro	Thr	Val	Arg	Ala	Ser	Gly	Leu	Lys	Thr	Gly	Val	Val	Cys	Cys
225					230					235					240
Asp	Ala	Ala	Gly	Trp	Asp	Arg	Gln	Val	Ala	Tyr	Thr	Glu	Ala	Ile	Glu
				245					250					255	
Ala	Asp	Pro	Glu	Ala	Ala	Lys	Ala	Val	Arg	Thr	Val	Thr	Gly	His	Arg
			260					265					270		
Tyr	Ser	Gly	Pro	Thr	Thr	Val	Pro	Gln	Pro	Thr	Asp	Lys	Arg	Val	Trp
		275					280					285			
Met	Ser	Glu	Trp	Ser	Pro	Asp	Gly	Thr	Thr	Trp	Asn	Glu	Asn	Trp	Asp
	290					295					300				
Asp	Gly	Ser	Gly	Tyr	Asp	Gly	Leu	Thr	Val	Ala	Ala	Asp	Ile	Gln	Asn
305					310					315					320
Thr	Leu	Thr	Val	Gly	Asn	Ala	Asn	Ala	Tyr	Val	Tyr	Trp	Thr	Gly	Ala
				325					330					335	
Ser	Leu	Gly	Ala	Thr	Arg	Gly	Leu	Ile	Gln	Leu	Ala	Asn	Pro	Gly	Asp
			340					345					350		
Ser	Tyr	Arg	Val	Ser	Lys	Arg	Tyr	Trp	Ala	Leu	Ala	Ala	Phe	Ser	Arg
		355					360					365			
Phe	Ile	Arg	Pro	Asp	Ala	Val	Arg	Val	Pro	Val	Thr	Asn	Ala	Asp	Pro
	370					375					380				
Ala	Leu	Ser	Val	Thr	Ala	Phe	Arg	Asn	Thr	Asp	Gly	Ser	Arg	Val	Ile
385					390					395					400
Glu	Ile	Leu	Asn	Thr	Ala	Thr	Thr	Glu	Lys	Ser	Ala	Gln	Phe	Ala	Leu
				405					410					415	
Arg	Gly	Gly	His	Asp	Arg	His	Pro	Glu	Gly	Tyr	Val	Thr	Asp	Glu	Thr
			420					425					430		
Arg	Ser	Ile	Thr	Pro	Ala	His	Val	Ala	Ser	Ala	Arg	Gly	Thr	Thr	Leu
		435					440					445			

Lys Ala Thr Leu Ala Pro Arg Ala Leu Thr Thr Ile Val Leu Asp
 450 455 460

<210> 315
 <211> 1224
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 315
 atgcggaacg tctgtcgtaa accattgaca atcggactcg ctttaacact attattgccc 60
 atgggaatga cggcaacatc agcgaagaat gcagattcct atgcgaaaaa acctcacatc 120
 agcgcatatga atgccccaca attggatcaa cgctacaaaa acgagttcac gattggtgcg 180
 gcagtagaac cttatcaact acaaaatgaa aaagacgtac aaatgctaaa gcgccacttc 240
 aacagcattg ttgccgagaa cgtaaatgaaa ccgatcagca ttcaacctga ggaaggaaaa 300
 ttcaattttg aacaagcgga tcaaatgtg aagttcgcta aggcaaatgg catggatatt 360
 cgcttcata cactcgtttg gcacagccaa gtacctcaat ggttctttct tgacaaggaa 420
 ggcaagccaa tggtaaatga aacagatcca gtgaaacgtg aacaaaataa acaactgctg 480
 ttaaaacgac ttgaaactca tattaacacg atcgtcgagc ggtacaaaaga tgacattaag 540
 tactgggacg ttgtaaatga ggttgtgggg gacgacggaa aactgcgcaa ctctccatgg 600
 tatcaaatcg ccggcatcga ttatatataa gtggcattcc aaacagcgag aaaatatggc 660
 ggcaacaaga ttaaaactta tatcaatgat tacaataccg aagtgggaacc aaagcgaagc 720
 gctctttata acttgggtgaa gcaattaaaa gaagagggcg ttcctattga cggcatcggc 780
 catcaatccc acattcaaat cggctggcct tctgaagcag aaatcgagaa aacgattaac 840
 atgttcgccc ctctcggtt agacaaccaa atcactgagc ttgatgtgag catgtacggc 900
 tggccgcccgc gcgcttaccg gacgtatgac gccattccaa aacaaaagt tttggatcag 960
 gcagcgcgct atgatcggtt gttcaaacgt tatgaaaagt tgagcgataa aattagcaac 1020
 gtacacttct ggggcatcgc cgacaatcat acgtggctcg acagccgtgc ggatgtgtac 1080
 tatgacgcca acgggaatgt tgtggttgac ccgaacgctc cgtacgcaaa agtggaaaaa 1140
 gggaaaggaa aagatgcgcc gttcgttttt ggaccggatt acaaaagtcaa acccgcatat 1200
 tgggctatta tcgaccacaa atag 1224

<210> 316
 <211> 407
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(28)

<400> 316
 Met Arg Asn Val Val Arg Lys Pro Leu Thr Ile Gly Leu Ala Leu Thr
 1 5 10 15
 Leu Leu Leu Pro Met Gly Met Thr Ala Thr Ser Ala Lys Asn Ala Asp
 20 25 30
 Ser Tyr Ala Lys Lys Pro His Ile Ser Ala Leu Asn Ala Pro Gln Leu
 35 40 45
 Asp Gln Arg Tyr Lys Asn Glu Phe Thr Ile Gly Ala Ala Val Glu Pro
 50 55 60
 Tyr Gln Leu Gln Asn Glu Lys Asp Val Gln Met Leu Lys Arg His Phe
 65 70 75 80
 Asn Ser Ile Val Ala Glu Asn Val Met Lys Pro Ile Ser Ile Gln Pro
 85 90 95
 Glu Glu Gly Lys Phe Asn Phe Glu Gln Ala Asp Arg Ile Val Lys Phe
 100 105 110
 Ala Lys Ala Asn Gly Met Asp Ile Arg Phe His Thr Leu Val Trp His
 115 120 125
 Ser Gln Val Pro Gln Trp Phe Phe Leu Asp Lys Glu Gly Lys Pro Met
 130 135 140
 Val Asn Glu Thr Asp Pro Val Lys Arg Glu Gln Asn Lys Gln Leu Leu
 145 150 155 160
 Leu Lys Arg Leu Glu Thr His Ile Lys Thr Ile Val Glu Arg Tyr Lys
 165 170 175
 Asp Asp Ile Lys Tyr Trp Asp Val Val Asn Glu Val Val Gly Asp Asp
 Page 238

180 185 190
 Gly Lys Leu Arg Asn Ser Pro Trp Tyr Gln Ile Ala Gly Ile Asp Tyr
 195 200 205
 Ile Lys Val Ala Phe Gln Thr Ala Arg Lys Tyr Gly Gly Asn Lys Ile
 210 215 220
 Lys Leu Tyr Ile Asn Asp Tyr Asn Thr Glu Val Glu Pro Lys Arg Ser
 225 230 235 240
 Ala Leu Tyr Asn Leu Val Lys Gln Leu Lys Glu Glu Gly Val Pro Ile
 245 250 255
 Asp Gly Ile Gly His Gln Ser His Ile Gln Ile Gly Trp Pro Ser Glu
 260 265 270
 Ala Glu Ile Glu Lys Thr Ile Asn Met Phe Ala Ala Leu Gly Leu Asp
 275 280 285
 Asn Gln Ile Thr Glu Leu Asp Val Ser Met Tyr Gly Trp Pro Pro Arg
 290 295 300
 Ala Tyr Pro Thr Tyr Asp Ala Ile Pro Lys Gln Lys Phe Leu Asp Gln
 305 310 315 320
 Ala Ala Arg Tyr Asp Arg Leu Phe Lys Leu Tyr Glu Lys Leu Ser Asp
 325 330 335
 Lys Ile Ser Asn Val Thr Phe Trp Gly Ile Ala Asp Asn His Thr Trp
 340 345 350
 Leu Asp Ser Arg Ala Asp Val Tyr Tyr Asp Ala Asn Gly Asn Val Val
 355 360 365
 Val Asp Pro Asn Ala Pro Tyr Ala Lys Val Glu Lys Gly Lys Gly Lys
 370 375 380
 Asp Ala Pro Phe Val Phe Gly Pro Asp Tyr Lys Val Lys Pro Ala Tyr
 385 390 395 400
 Trp Ala Ile Ile Asp His Lys
 405

<210> 317
 <211> 1695
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 317
 gtggctggaa gctcgtcac gagcaacggc ctctcggcca ttctctcgtt ccagtcggac 60
 tggggcagcg gttactgctc gacggtagaa cttcagaacg tcggcggaac tccgatcacg 120
 gcgtgggagg tccaggtgga gctcgtctgg acgaccgtga acagcagcca cagcgcggcg 180
 ttctcctcga caggcaccgg cctggctcgcc aagcccttgt cctggaacgc gacgctggca 240
 cccgccgcca agacgacctt cggcttctgc gcggccgctc cgagcgcagc ggcgcgccc 300
 tccgtggtgc aagtgcagc gaacggctcc gccaccggaa cgggcggaac gagcggcggc 360
 ggcacgggcg gctcgcaccg tacgggcggc tcgaccgcta cgggcggctc cgtgggtcg 420
 accgcgggag tgtgcgcggc aacctacgag gccgagagca tgctccacag caccggcaac 480
 gccatcacgc gcggtggaa catctattcg aacggcaaca tcaccgccac gcactccttc 540
 gcagccggct cgaatcgact caccgtgcac gccaaaggcg accaggccaa cggggcgccc 600
 atcatgcgcg tcagcgtggg caacaccgtc gtcggcgagg tgccagtgc ggtgaccgtg 660
 tggacaccgt actgcttcca ctacgcccg gcgagcgcag gcgcgcagac cgtcaagatc 720
 gagttcacga acgactacaa tggcggcacc ggcggcacc gcaatctgca cgtggacaag 780
 gtcgcggtgc agtgcggcgc gagctgcaac agcgggagcg gagggggcac cggcggctcg 840
 agcgggaagc gcggcacctc ggccaccggc ggctccgcca gcggtggcgc ggcagggacg 900
 acctgcacga acgttcgtcc cactggaacc gactgggagc cggcgacctg cgacatgtg 960
 gcctcgcaaa ctagcagtg cagcgcggcc tggatgatcg acaaccatta ctgcgaccag 1020
 agctgcgggc ggtgctcggg cgggagcggg accggtggca cgaacacggg aggcaccggc 1080
 ggtggagtga ccccgagtac ctgcacggag cccaattctc agcagtgtc cactacaag 1140
 gtcgggactc actgcggcct cacctacgag atctggaccg acggctccgc gggctgcatg 1200
 acgaacacct cctacgggtt cctcgccaat tggagccagg ggaacgcaaa ctacctggct 1260
 cgcaagggcg ttcggccccg ctctcgcga ccggtctgta cgtacagcgc gaactaccag 1320
 ccgaacggga attcctacct ggggatctac ggttggacgc agaaccgct cgtcgagtac 1380
 tacatcatcg atagctgggg gagctggcgt ccaccgggga cccaggcgat gggcaccgtc 1440
 caggtggacg gcgggacctc cgatatctac cggagcgagc gggtgaaaca gccctcgatc 1500
 gagggcaaca agaccttctg gcagtactgg agcgtccgca cccagaagcg caccagtggg 1560
 accatcacg tggctccgca cttcgccgct tgggcggcat ccggactgca gatgggctcc 1620
 ttctacgagg ttccctgggt ggtggagggc tacaacagct ccggcagcgc cgacgtaacg 1680
 gtgtcgttcc ggtag 1695

<210> 318
 <211> 564
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 318

```

Met Ala Gly Ser Ser Leu Thr Ser Asn Gly Leu Ser Ala Ile Leu Ser
 1      5      10      15
Leu Gln Ser Asp Trp Gly Ser Gly Tyr Cys Ala Thr Val Glu Leu Gln
 20      25      30
Asn Val Gly Gly Thr Pro Ile Thr Ala Trp Glu Val Gln Val Glu Leu
 35      40      45
Ala Gly Thr Thr Val Asn Ser Ser His Ser Ala Ala Phe Ser Ser Thr
 50      55      60
Gly Thr Arg Leu Val Ala Lys Pro Leu Ser Trp Asn Ala Thr Leu Ala
 65      70      75
Pro Ala Ala Lys Thr Phe Gly Phe Cys Ala Ala Ala Pro Ser Ala
 85      90      95
Ala Ala Arg Pro Ser Val Val Gln Val Thr Ala Asn Gly Ser Ala Thr
100      105      110
Gly Thr Gly Thr Ser Gly Gly Gly Thr Gly Gly Ser Thr Ala Thr
115      120      125
Gly Gly Ser Thr Ala Thr Gly Gly Ser Gly Gly Ser Thr Ala Gly Val
130      135      140
Cys Ala Ala Thr Tyr Glu Ala Glu Ser Met Leu His Ser Thr Gly Asn
145      150      155
Ala Ile Ser Gly Gly Trp Asn Ile Tyr Ser Asn Gly Asn Ile Thr Ala
165      170      175
Thr His Ser Phe Ala Ala Gly Ser Asn Arg Leu Thr Val His Ala Lys
180      185      190
Gly Asp Gln Ala Asn Gly Ala Pro Ile Met Arg Val Ser Val Gly Asn
195      200      205
Thr Val Val Gly Glu Val Pro Val Pro Val Thr Val Trp Thr Pro Tyr
210      215      220
Cys Phe Asp Tyr Ala Ala Ser Ala Gly Ala Gln Thr Val Lys Ile
225      230      235
Glu Phe Thr Asn Asp Tyr Asn Gly Gly Thr Gly Ala Asp Arg Asn Leu
245      250      255
His Val Asp Lys Val Ala Val Gln Cys Gly Ala Ser Cys Asn Ser Gly
260      265      270
Ser Gly Gly Thr Gly Gly Ser Ser Gly Ser Gly Gly Thr Ser Ala
275      280      285
Thr Gly Gly Ser Ala Ser Gly Gly Ala Ala Gly Thr Thr Cys Thr Asn
290      295      300
Val Arg Pro Thr Gly Thr Asp Trp Asp Ala Ala Thr Cys Asp Met Trp
305      310      315
Ala Ser Gln Thr Ser Glu Cys Ser Ala Ala Trp Met Ile Asp Asn His
325      330      335
Tyr Cys Asp Gln Ser Cys Gly Arg Cys Ser Gly Gly Ser Gly Thr Gly
340      345      350
Gly Thr Asn Thr Gly Gly Thr Gly Gly Gly Val Thr Pro Ser Thr Cys
355      360      365
Thr Glu Pro Asn Ser Gln Gln Cys Ser Thr Tyr Lys Val Gly Thr His
370      375      380
Cys Gly Leu Thr Tyr Glu Ile Trp Thr Asp Gly Ser Ala Gly Cys Met
385      390      395
Thr Asn Thr Ser Tyr Gly Phe Leu Ala Asn Trp Ser Gln Gly Asn Ala
405      410      415
Asn Tyr Leu Ala Arg Lys Gly Val Arg Pro Gly Ser Ser Arg Pro Val
420      425      430
Val Thr Tyr Ser Ala Asn Tyr Gln Pro Asn Gly Asn Ser Tyr Leu Gly
435      440      445
Ile Tyr Gly Trp Thr Gln Asn Pro Leu Val Glu Tyr Tyr Ile Ile Asp
450      455      460
Ser Trp Gly Ser Trp Arg Pro Pro Gly Thr Gln Ala Met Gly Thr Val
465      470      475      480

```

Gln Val Asp Gly Gly Thr Tyr Asp Ile Tyr Arg Ser Glu Arg Val Asn
 485 490 495
 Lys Pro Ser Ile Glu Gly Asn Lys Thr Phe Trp Gln Tyr Trp Ser Val
 500 505 510
 Arg Thr Gln Lys Arg Thr Ser Gly Thr Ile Thr Val Ala Pro His Phe
 515 520 525
 Ala Ala Trp Ala Ala Ser Gly Leu Gln Met Gly Ser Phe Tyr Glu Val
 530 535 540
 Ser Leu Val Val Glu Gly Tyr Asn Ser Ser Gly Ser Ala Asp Val Thr
 545 550 555 560
 Val Ser Phe Arg

<210> 319
 <211> 1095
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 319
 atgaaggtga cccgaacagc tgtcgcgggc attgtcgcgc cagcggtcct catcacgatc 60
 ggcacgtcga ccgcgtcggc tgaggatgaa ccaaccagcg agaacacgtc gacggatcag 120
 ccgttgcgcg tcctggcagc caaagccggg atcgcgttcg gcacggccgt cgacatgaac 180
 gcgtacaaca acgacgcgac ctaccgtgag ctcgctcgcc aggagtctc gagcgtcacg 240
 gccgagaacg tcatgaagtg gcagctcctc gagccgcagc gaggggtcta caactgggggt 300
 ccggccgatc agctcgtgcg cgtagccaac gagaacggcc agaagggtgcg cgggcacacg 360
 ctcatctggc acaaccagct gcccacctgg cttaccagcg gattcgcctc cgggtgagatc 420
 acaccggacg agctccggca gctcctgagg aaccacatct tcacggtgat gcgccacttc 480
 aagggcgaga tccaccagtg ggatgtcgcc aacgaggtca tcgacgacag cggcaacctg 540
 cgcaacacga tctggctgca gaacctgggt ccgagctaca tcgcgacgc gttccggtgg 600
 gctcgcaagg ccgacccgga cgccgccctc tatctgaacg actacaacgt cgagggcccg 660
 aacgccaagg ccgatgcgta ctacgccctg gtcaagcagc tcctcgccga cgacgtgccg 720
 gtggacggct tcggaataca ggggcacctc ggtgtgcagt tcggcttctg gcccgcgagt 780
 gcggtggccg acaacatggg gcgcttcgag gcactcggcc tgcagacggc ggtcaccgag 840
 gcggatgtcc ggatgatcat gccgcccgcg gaggacaagc tggccgcaca ggcacgtggc 900
 tacagcacgt tggctccagg ctgcctgatg gccaaagcgt gcaggtcgtt caccgtctgg 960
 ggcttcaccg acaagtactc ctgggttcgc ggcaaccttc ccggccaggc cgcggcgaac 1020
 ctctggccg aggacttcca gcccaagccg gcttactacg ccgtccagga tgacctcgcg 1080
 cgcgcccggac ggtag 1095

<210> 320
 <211> 364
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(27)

<400> 320
 Met Lys Val Thr Arg Thr Ala Val Ala Gly Ile Val Ala Ala Ala Val
 1 5 10 15
 Leu Ile Thr Ile Gly Thr Ser Thr Ala Ser Ala Glu Asp Glu Pro Thr
 20 25 30
 Ser Glu Asn Thr Ser Thr Asp Gln Pro Leu Arg Val Leu Ala Ala Lys
 35 40 45
 Ala Gly Ile Ala Phe Gly Thr Ala Val Asp Met Asn Ala Tyr Asn Asn
 50 55 60
 Asp Ala Thr Tyr Arg Glu Leu Val Gly Gln Glu Phe Ser Ser Val Thr
 65 70 75 80
 Ala Glu Asn Val Met Lys Trp Gln Leu Leu Glu Pro Gln Arg Gly Val
 85 90 95
 Tyr Asn Trp Gly Pro Ala Asp Gln Leu Val Arg Val Ala Asn Glu Asn
 100 105 110
 Gly Gln Lys Val Arg Gly His Thr Leu Ile Trp His Asn Gln Leu Pro

Thr	Trp	Leu	Thr	Ser	Gly	Val	Ala	Ser	Gly	Glu	Ile	Thr	Pro	Asp	Glu
130						135					140				
Leu	Arg	Gln	Leu	Leu	Arg	Asn	His	Ile	Phe	Thr	Val	Met	Arg	His	Phe
145					150					155					160
Lys	Gly	Glu	Ile	His	Gln	Trp	Asp	Val	Ala	Asn	Glu	Val	Ile	Asp	Asp
				165					170					175	
Ser	Gly	Asn	Leu	Arg	Asn	Thr	Ile	Trp	Leu	Gln	Asn	Leu	Gly	Pro	Ser
			180					185					190		
Tyr	Ile	Ala	Asp	Ala	Phe	Arg	Trp	Ala	Arg	Lys	Ala	Asp	Pro	Asp	Ala
		195					200					205			
Ala	Leu	Tyr	Leu	Asn	Asp	Tyr	Asn	Val	Glu	Gly	Pro	Asn	Ala	Lys	Ala
	210					215					220				
Asp	Ala	Tyr	Tyr	Ala	Leu	Val	Lys	Gln	Leu	Leu	Ala	Asp	Asp	Val	Pro
225					230					235					240
Val	Asp	Gly	Phe	Gly	Ile	Gln	Gly	His	Leu	Gly	Val	Gln	Phe	Gly	Phe
				245					250					255	
Trp	Pro	Ala	Ser	Ala	Val	Ala	Asp	Asn	Met	Gly	Arg	Phe	Glu	Ala	Leu
			260					265					270		
Gly	Leu	Gln	Thr	Ala	Val	Thr	Glu	Ala	Asp	Val	Arg	Met	Ile	Met	Pro
		275					280					285			
Pro	Asp	Glu	Asp	Lys	Leu	Ala	Ala	Gln	Ala	Arg	Gly	Tyr	Ser	Thr	Leu
	290					295					300				
Val	Gln	Gly	Cys	Leu	Met	Ala	Lys	Arg	Cys	Arg	Ser	Phe	Thr	Val	Trp
305					310					315					320
Gly	Phe	Thr	Asp	Lys	Tyr	Ser	Trp	Val	Pro	Gly	Thr	Phe	Pro	Gly	Gln
				325					330					335	
Gly	Ala	Ala	Asn	Leu	Leu	Ala	Glu	Asp	Phe	Gln	Pro	Lys	Pro	Ala	Tyr
			340					345					350		
Tyr	Ala	Val	Gln	Asp	Asp	Leu	Ala	Arg	Ala	Gly	Arg				
		355					360								

<210> 321

<211> 1608

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 321

gtggactggg	gggacgtgga	tattttttcc	gccaaggaaa	tcaccacccc	gcaactggca	60
accttcctgg	atgcctcacg	agaccatcgc	aagccgggtca	tgatcggcga	gatgacccca	120
cgccacgtcg	gcgtgatcga	ggggcagaaa	tgctgggatg	aatgggtttg	cccgatgatt	180
gatctgctca	aacgtcgccc	cgaaatcaag	gccacggcct	atatcaactg	ggaatggcgc	240
gagtggtccg	accgcctcgg	cttcgcgtgg	cacaactggg	gcgacgcccg	catcgagggc	300
aacgcccttg	ttcgtgatcg	ctgggtgcag	gaactctccc	accccatcta	tctccacgcg	360
gcgcgcgacg	gatcttgtcc	gctgcccga	atcacgcgcc	tcccatccgc	gaccccgicg	420
ctccagaccg	tgttccagga	ccatttctcg	atgggtgctg	ccttgaatgt	gaggcagttc	480
accgaaaacg	acgcaaccaa	gaccgctctc	atcaagaagc	aattcaacac	catcacgccc	540
gagaatgttc	tcaagtgggg	gccggttcat	cctgagccca	accggttcaa	cttcgaatcc	600
accgatcggt	acgtggactt	tggtgtgaag	aaccggatgt	tcacggtcgg	ccacaccctc	660
gtctggcacc	accagacacc	cgctgggtg	tttcaagatt	cccaaggcca	gccgctcgac	720
cggtatggac	tgctcaatcg	cttgagcaac	cacatccaca	cggtgggttg	acgctacaag	780
ggccgcaccc	acgggtggga	tatggtgaac	gaggccttga	acgatgacgg	caccctccgc	840
cctagccaat	ggcttaaaat	catcgcccc	gactacattg	caaagcggtt	tgcccttgcc	900
cacgcgcgcg	accctgcccg	tgaactgtat	tacaacgatt	acagtctcga	tcattcccgc	960
aagtgtgctg	gtgcatcg	gctgggtgaag	cagctccaga	cgaatggcat	atccattgcc	1020
gggatggca	cgcagaccca	cgctcgactc	aacggacctt	ccccccagtc	ggtggatgat	1080
tcattgacgg	cctttggcca	gctcgcgctg	aaggtcatgg	ttaccgaact	cgacgttgat	1140
gtgctgcccg	cgcgcagcca	aaatcaaaac	gcggatctca	accagcccgc	cttgtccaat	1200
ccgcgccctca	atcccgccct	caatccctat	cccgatgggc	tgccgcaagc	cgctcaggac	1260
aaactggccg	ctcgctatgc	ggaactcttc	gccgtgttcg	tcaagcacgc	cgacaaaatc	1320
agccgcgtca	cgttctggtg	cgtcaccgac	ggcgactcct	ggctgaacaa	ctggcccgtg	1380
cgtggccgcg	tcaactatcc	gctgctgttc	gaccgtgccca	gccagcccaa	gcccgccttc	1440
gatgcgggtca	ttcgcgtcgc	caaggacccc	ccgacgggtt	cgacacatct	caccccgctc	1500
cacgatgcgg	cgccgggtcct	ggtcaatccg	cacaagggct	ggtaccacca	ctacccggac	1560
aatcacatca	acaagtatga	gatcgccccg	gatgccgacc	tgacggaa		1608

<210> 322
 <211> 536
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 322

```

Met Asp Trp Trp Asp Val Asp Ile Phe Ser Ala Lys Glu Ile Thr His
 1      5      10      15
Pro Gln Leu Ala Thr Phe Leu Asp Ala Ser Arg Asp His Arg Lys Pro
      20      25      30
Val Met Ile Gly Glu Met Thr Pro Arg His Val Gly Val Ile Glu Gly
      35      40      45
Gln Lys Cys Trp Asp Glu Trp Phe Gly Pro Met Ile Asp Leu Leu Lys
      50      55      60
Arg Arg Pro Glu Ile Lys Ala Thr Ala Tyr Ile Asn Trp Glu Trp Arg
      65      70      75      80
Glu Trp Ser Asp Arg Leu Gly Phe Arg Trp His Asn Trp Gly Asp Ala
      85      90      95
Arg Ile Glu Gly Asn Ala Leu Val Arg Asp Arg Trp Val Gln Glu Leu
      100      105      110
Ser His Pro Ile Tyr Leu His Ala Ala Arg Asp Gly Ser Cys Pro Leu
      115      120      125
Pro Pro Ile Thr Ala Leu Pro Ser Ala Thr Pro Ser Leu Gln Thr Val
      130      135      140
Phe Gln Asp His Phe Leu Met Gly Ala Ala Leu Asn Val Arg Gln Phe
      145      150      155      160
Thr Glu Asn Asp Ala Thr Lys Thr Ala Leu Ile Lys Lys Gln Phe Asn
      165      170      175
Thr Ile Thr Pro Glu Asn Val Leu Lys Trp Gly Pro Val His Pro Glu
      180      185      190
Pro Asn Arg Phe Asn Phe Glu Ser Thr Asp Arg Tyr Val Asp Phe Gly
      195      200      205
Val Lys Asn Arg Met Phe Ile Val Gly His Thr Leu Val Trp His His
      210      215      220
Gln Thr Pro Ala Trp Val Phe Gln Asp Ser Gln Gly Gln Pro Leu Asp
      225      230      235      240
Arg Asp Gly Leu Leu Asn Arg Leu Ser Asn His Ile His Thr Val Val
      245      250      255
Gly Arg Tyr Lys Gly Arg Ile His Gly Trp Asp Met Val Asn Glu Ala
      260      265      270
Leu Asn Asp Asp Gly Thr Leu Arg Pro Ser Gln Trp Leu Lys Ile Ile
      275      280      285
Gly Pro Asp Tyr Ile Ala Lys Ala Phe Ala Leu Ala His Ala Ala Asp
      290      295      300
Pro Ala Ala Glu Leu Tyr Tyr Asn Asp Tyr Ser Leu Asp His Pro Ala
      305      310      315      320
Lys Cys Ala Gly Ala Ile Ala Leu Val Lys Gln Leu Gln Thr Asn Gly
      325      330      335
Ile ser Ile Ala Gly Ile Gly Thr Gln Thr His Val Gly Leu Asn Gly
      340      345      350
Pro ser Pro Gln Ser Val Asp Asp Ser Leu Thr Ala Phe Gly Gln Leu
      355      360      365
Gly Val Lys Val Met Val Thr Glu Leu Asp Val Asp Val Leu Pro Ala
      370      375      380
Ala Ser Gln Asn Gln Asn Ala Asp Leu Asn Gln Pro Ala Leu Ser Asn
      385      390      395      400
Pro Ala Leu Asn Pro Ala Leu Asn Pro Tyr Pro Asp Gly Leu Pro Gln
      405      410      415
Ala Val Gln Asp Lys Leu Ala Ala Arg Tyr Ala Glu Leu Phe Ala Val
      420      425      430
Phe Val Lys His Ala Asp Lys Ile Ser Arg Val Thr Phe Trp Cys Val
      435      440      445
Thr Asp Gly Asp Ser Trp Leu Asn Asn Trp Pro Val Arg Gly Arg Val
      450      455      460
Asn Tyr Pro Leu Leu Phe Asp Arg Ala Ser Gln Pro Lys Pro Ala Phe
      465      470      475      480

```

Asp Ala Val Ile Arg Val Ala Lys Asp Pro Pro Thr Val Ser His Asn
 485 490 495
 Leu Thr Pro Leu His Asp Ala Ala Arg Val Leu Val Asn Pro His Lys
 500 505 510
 Gly Trp Tyr His His Tyr Pro Asp Asn His Ile Asn Lys Tyr Glu Ile
 515 520 525
 Ala Arg Asp Ala Asp Leu Thr Glu
 530 535

<210> 323

<211> 2355

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 323

atgatgctca	atgcccgttg	tatccaactt	atgaagttgt	tgcttcgctc	ttctctttat	60
cttacccgtg	acaaattggc	gcaatcattg	aatgtatcca	agcgaacgat	ttattacgat	120
atacaaaaaa	cgaatgaatg	gttgcattcat	gaagggtcga	agccgattca	atatgctgctc	180
gggctcggat	ttcgcttgga	tgatgaagtg	aaacaagaaa	taacaacaaa	gtggaacaca	240
ttacaacctg	cccgcacatta	cacatatcag	tcatgggagc	gaaaagcttg	gattgggtta	300
tggattttga	ctcgcgttca	tccactgtat	ttgtctgatt	ttttagagaa	attacatgta	360
agcaggagca	cgttgttaaa	tgacataaag	gaactgaaag	aagattggca	gtcatttcag	420
ttgcgattgt	cattccatcg	caaaaaaggg	tatttttcat	caggggaaga	aatccaaaaa	480
aggaaattga	tgattcgtta	tattcatcaa	atattagcgg	cgatggatga	ccagcatttc	540
gctgcagaat	tgtcagctga	gtgtcaatgg	ccaatctttg	attggatttg	ccaattcgag	600
tctacttttt	ctattcgccta	taccggtgag	gttattcaaa	ctttacctat	ttacctcgca	660
ttgttccaaa	gcagggtggc	tagaggcaaa	tttgtgcaaa	tggaacgagca	agaaaaagaa	720
gtgctaaggt	caatgcggga	ataccagatt	gctgatcatc	tcgttagacg	aattgaaaac	780
gtttccgaaa	tatctattcc	cgatgacgag	gtttgttatt	tgacgaccca	tttactcagt	840
tttcgagttg	cagatgacaa	gcaaatcgat	cataacgatg	acatcactac	tttgaacga	900
atcattccgac	atatggtgga	tgattttcaa	acttatgcct	gtgtacaatt	caagcgtcgc	960
gaagagttgg	aaaaaaaggt	attggttcat	atgaagcctg	cctattatcg	actgaaatac	1020
ggttttcatc	tgcaaaacga	tctgaccgaa	tcggtcaaa	cgaactatca	agatttattt	1080
accttaacga	aaaaagtcgt	ccatcatitta	gaaagtgtag	tgggccagcc	ggtcagcgac	1140
gatgaaattg	cttatatcgc	catgcatttt	ggcggatggg	tggaacagaga	gggggtgtcg	1200
gttccagtac	ggaaaaaggt	gttgatcgtc	tgcgagagcg	ggattggaac	atcgcgaaatg	1260
ttgcaaaaac	aattggatca	acgctacaaa	aacgagttca	cgattggtgc	ggcagtagaa	1320
ccttatcaac	tacaaaatga	aaaagacgta	caaagtctaa	agcgccactt	caacagcatt	1380
gttgccgaga	acgtaatgaa	accgatcagc	attcaacctg	aggaaggaaa	attcaatttt	1440
gaacaagcgg	atcgaattgt	gaagttcgct	aaggcaaatg	gcatggatat	tcgcttccat	1500
acactcgttt	ggcacagcca	agtacctcaa	tggttctttc	ttgacaagga	aggcaagcca	1560
atggttaatg	aaacagatcc	agtgaacgtg	gaacaaaata	aacaactgct	gttaaaacga	1620
cttgaaactc	atattaaaac	gatcgctcag	cggtacaaag	atgacattaa	gtactgggac	1680
gttgtaaattg	aggttgtggg	ggacgacgga	aaactgcgca	actctccatg	gtatcaaatc	1740
gccggcatcg	attatattaa	agtggcattc	caaacagcga	gaaaatatgg	cggcaacaag	1800
attaaacttt	atatcaatga	ttacaatacc	gaagtggaa	caaagcgaag	cgctctttat	1860
aacttggtga	agcaattaaa	agaagagggc	gttcctattg	acggcatcgg	ccatcaatcc	1920
cacattcaaa	tcggctggcc	ttctgaagca	gaaatcgaga	aaacgattaa	catgttcgcc	1980
gctctcggct	tagacaacca	aatcactgag	cttgatgtga	gcatgtacgg	ttggccgccc	2040
cgcgcttacc	cgacgtatga	cgcatttcca	aaacaaaagt	ttttggatca	ggcagcgctc	2100
tatgatcggt	tgttcaaaact	gtatgaaaag	ttgagcgata	aaattagcaa	cgtcaccttc	2160
tggggcatcg	ccgacaatca	tacgtggctc	gacagccgtg	cggtatgtgt	ctatgacgcc	2220
aacgggaatg	ttgtggttga	cccgaacgct	ccgtacgcaa	aagtggaaaa	agggaaagga	2280
aaagatgcgc	cgttcgtttt	tggaccggat	tacaaagtca	aaccgcgata	ttgggctatt	2340
atcgaccaca	aatag					2355

<210> 324

<211> 784

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 324

Met Met Leu Asn Ala Arg Cys Ile Gln Leu Met Lys Leu Leu Leu Arg
 Page 244

1				5					10				15
Ser	Ser	Leu	Tyr	Leu	Thr	Ala	Asp	Lys	Leu	Ala	Gln	Ser	Leu
			20					25					30
Ser	Lys	Arg	Thr	Ile	Tyr	Tyr	Asp	Ile	Gln	Lys	Thr	Asn	Glu
		35					40					45	Trp
His	His	Glu	Gly	Leu	Lys	Pro	Ile	Gln	Tyr	Ala	Arg	Gly	Leu
	50					55				60			Gly
Arg	Leu	Asp	Asp	Glu	Val	Lys	Gln	Glu	Ile	Thr	Thr	Lys	Trp
	65				70					75			Asn
Leu	Gln	Pro	Ala	Arg	His	Tyr	Thr	Tyr	Gln	Ser	Trp	Glu	Arg
				85					90				Lys
Trp	Ile	Gly	Leu	Trp	Ile	Leu	Thr	Arg	Val	His	Pro	Leu	Tyr
		100						105				110	Leu
Asp	Phe	Leu	Glu	Lys	Leu	His	Val	Ser	Arg	Ser	Thr	Leu	Leu
		115					120					125	Asn
Ile	Lys	Glu	Leu	Lys	Glu	Asp	Trp	Gln	Ser	Phe	Gln	Leu	Arg
	130					135					140		Leu
Phe	His	Arg	Lys	Lys	Gly	Tyr	Phe	Ser	Ser	Gly	Glu	Glu	Ile
	145				150					155			Gln
Arg	Lys	Leu	Met	Ile	Arg	Tyr	Ile	His	Gln	Ile	Leu	Ala	Ala
				165					170				Met
Asp	Gln	His	Phe	Ala	Ala	Glu	Leu	Ser	Ala	Glu	Cys	Gln	Trp
			180					185					Pro
Phe	Asp	Trp	Ile	Cys	Gln	Phe	Glu	Ser	Thr	Phe	Ser	Ile	Arg
		195					200					205	Tyr
Gly	Glu	Val	Ile	Gln	Thr	Leu	Pro	Ile	Tyr	Leu	Ala	Leu	Phe
	210					215					220		Gln
Arg	Trp	Ala	Arg	Gly	Lys	Phe	Val	Gln	Met	Asp	Glu	Gln	Glu
	225				230					235			Lys
Val	Leu	Arg	Ser	Met	Arg	Glu	Tyr	Gln	Ile	Ala	Asp	His	Leu
				245					250				Val
Arg	Ile	Glu	Asn	Val	Ser	Glu	Ile	Ser	Ile	Pro	Asp	Asp	Glu
			260					265					Val
Tyr	Leu	Thr	Thr	His	Leu	Leu	Ser	Phe	Arg	Val	Ala	Asp	Asp
		275					280					285	Lys
Ile	Asp	His	Asn	Asp	Asp	Ile	Thr	Thr	Leu	Lys	Arg	Ile	Ile
	290					295					300		Arg
Met	Val	Asp	Asp	Phe	Gln	Thr	Tyr	Ala	Cys	Val	Gln	Phe	Lys
	305				310					315			Arg
Glu	Glu	Leu	Glu	Lys	Asn	Leu	Leu	Val	His	Met	Lys	Pro	Ala
				325					330				Tyr
Arg	Leu	Lys	Tyr	Gly	Phe	His	Leu	Gln	Asn	Asp	Leu	Thr	Glu
			340					345				350	Ser
Lys	Ala	Asn	Tyr	Gln	Asp	Leu	Phe	Thr	Leu	Thr	Lys	Lys	Val
		355					360					365	His
His	Leu	Glu	Ser	Val	Val	Gly	Gln	Pro	Val	Ser	Asp	Asp	Glu
	370					375					380		Ile
Tyr	Ile	Ala	Met	His	Phe	Gly	Gly	Trp	Leu	Asp	Arg	Glu	Gly
	385				390					395			Val
Val	Pro	Val	Arg	Lys	Lys	Val	Leu	Ile	Val	Cys	Glu	Ser	Gly
				405					410				Ile
Thr	Ser	Arg	Met	Leu	Gln	Lys	Gln	Leu	Asp	Gln	Arg	Tyr	Lys
			420					425				430	Asn
Phe	Thr	Ile	Gly	Ala	Ala	Val	Glu	Pro	Tyr	Gln	Leu	Gln	Asn
		435					440					445	Glu
Asp	Val	Gln	Met	Leu	Lys	Arg	His	Phe	Asn	Ser	Ile	Val	Ala
	450					455					460		Glu
Val	Met	Lys	Pro	Ile	Ser	Ile	Gln	Pro	Glu	Glu	Gly	Lys	Phe
	465				470					475			Asn
Glu	Gln	Ala	Asp	Arg	Ile	Val	Lys	Phe	Ala	Lys	Ala	Asn	Gly
				485					490				Met
Ile	Arg	Phe	His	Thr	Leu	Val	Trp	His	Ser	Gln	Val	Pro	Gln
			500					505				510	Trp
Phe	Leu	Asp	Lys	Glu	Gly	Lys	Pro	Met	Val	Asn	Glu	Thr	Asp
		515					520					525	Pro
Lys	Arg	Glu	Gln	Asn	Lys	Gln	Leu	Leu	Leu	Lys	Arg	Glu	Thr
	530					535					540		His
Ile	Lys	Thr	Ile	Val	Glu	Arg	Tyr	Lys	Asp	Asp	Ile	Lys	Tyr
	545				550				555				Trp
													Asp
													560

Val Val Asn Glu Val Val Gly Asp Asp Gly Lys Leu Arg Asn Ser Pro
 Trp Tyr Gln Ile Ala Gly Ile Asp Tyr Ile Lys Val Ala Phe Gln Thr
 Ala Arg Lys Tyr Gly Gly Asn Lys Ile Lys Leu Tyr Ile Asn Asp Tyr
 Asn Thr Glu Val Glu Pro Lys Arg Ser Ala Leu Tyr Asn Leu Val Lys
 Gln Leu Lys Glu Glu Gly Val Pro Ile Asp Gly Ile Gly His Gln Ser
 His Ile Gln Ile Gly Trp Pro Ser Glu Ala Glu Ile Glu Lys Thr Ile
 Asn Met Phe Ala Ala Leu Gly Leu Asp Asn Gln Ile Thr Glu Leu Asp
 Val Ser Met Tyr Gly Trp Pro Pro Arg Ala Tyr Pro Thr Tyr Asp Ala
 Ile Pro Lys Gln Lys Phe Leu Asp Gln Ala Ala Arg Tyr Asp Arg Leu
 Phe Lys Leu Tyr Glu Lys Leu Ser Asp Lys Ile Ser Asn Val Thr Phe
 Trp Gly Ile Ala Asp Asn His Thr Trp Leu Asp Ser Arg Ala Asp Val
 Tyr Tyr Asp Ala Asn Gly Asn Val Val Val Asp Pro Asn Ala Pro Tyr
 Ala Lys Val Glu Lys Gly Lys Gly Lys Asp Ala Pro Phe Val Phe Gly
 Pro Asp Tyr Lys Val Lys Pro Ala Tyr Trp Ala Ile Ile Asp His Lys

<210> 325
 <211> 1146
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 325
 atgactattt cccgccggaa atttatgtgg ggcacagctg cactcctggc caccaccag 60
 ctcaaaaccc gcgctctcgc cgctgccatg gccagcacag gcatcaagga cgccttcaag 120
 ggcgacttcc atatcggcac cgccatcagc aacgctaccc tgcaaaacca ggatgccacc 180
 atgctggatt tgatcaagcg cgaatttaat gcaattaccg ctgaaaattg catgaagtgg 240
 gagcctattc gccacagct ggatcagtgg aattgggagc tggccgaccg ctttgtggat 300
 ttcggcggtta aaaacaagat gtatgtggta ggtcacacgc tgatttggca cagccaggcg 360
 ccagcgcaca tttatctcga cgccgatggt aagcccaaca gtcgcgatgc ccagttgaaa 420
 gtaatggagg agcacatacg taccctggcg ggccgctaca aaggaaagat agacgcctgg 480
 gacgtgggta acgaagcagt ggaggatgat ggcagctggc gtcaaaccgg ctggtacaaa 540
 aacatgggtg aagaatatat cgcccatgcc ttccgcttgg cagccgaggt agaccccaac 600
 gccaagctac tctacaacga ctacaacgag gctgtaccgg ccaagcgtga tgcgattatt 660
 cggttggtta aaggcgtgca gaaggctggc gcacccattc acgggtgtgg gatgcaaggg 720
 cacatgagcc tgtcacatcc ggatttcgcg gagttcgaaa aatccataat cgaatacgcc 780
 aagttggggg tgaaggtgca cgttaccgaa ctggatatcg acgtgttgcc actggcgtgg 840
 aacctgagtg cggaaatttc caatcgcttt gaataccgcc cagagatgga tccttatcgc 900
 gaaggtttgc ccgcaaagt cgaggaggag ctagcggctc gttacgaggc gctgtttaaa 960
 atcctgctgc gtcatcgcg caaaattgag cgtgtgacca cttggggcac caacgactca 1020
 gagacctggg taaatggctt ccccatccg gggcgcatga attacccaat gctgttcgat 1080
 cgtaataacc agcccaagtt ggcctatcac cggctgctgg cactcaaaca aaagaaaagt 1140
 cagtaa 1146

<210> 326
 <211> 381
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(27)

<400> 326
Met Thr Ile Ser Arg Arg Lys Phe Met Trp Gly Thr Ala Ala Leu Leu
1 5 10 15
Ala Thr Thr Gln Leu Lys Thr Arg Ala Leu Ala Ala Met Ala Ser
20 25 30
Thr Gly Ile Lys Asp Ala Phe Lys Gly Asp Phe His Ile Gly Thr Ala
35 40 45
Ile Ser Asn Ala Thr Leu Gln Asn Gln Asp Ala Thr Met Leu Asp Leu
50 55 60
Ile Lys Arg Glu Phe Asn Ala Ile Thr Ala Glu Asn Cys Met Lys Trp
65 70 75 80
Glu Pro Ile Arg Pro Gln Leu Asp Gln Trp Asn Trp Glu Leu Ala Asp
85 90 95
Arg Phe Val Asp Phe Gly Val Lys Asn Lys Met Tyr Val Val Gly His
100 105 110
Thr Leu Ile Trp His Ser Gln Ala Pro Ala His Ile Tyr Leu Asp Ala
115 120 125
Asp Gly Lys Pro Asn Ser Arg Asp Ala Gln Leu Lys Val Met Glu Glu
130 135 140
His Ile Arg Thr Leu Ala Gly Arg Tyr Lys Gly Lys Ile Asp Ala Trp
145 150 155 160
Asp Val Val Asn Glu Ala Val Glu Asp Asp Gly Ser Trp Arg Gln Thr
165 170 175
Gly Trp Tyr Lys Asn Met Gly Glu Glu Tyr Ile Ala His Ala Phe Arg
180 185 190
Leu Ala Ala Glu Val Asp Pro Asn Ala Lys Leu Leu Tyr Asn Asp Tyr
195 200 205
Asn Glu Ala Val Pro Ala Lys Arg Asp Ala Ile Ile Arg Val Val Lys
210 215 220
Gly Val Gln Lys Ala Gly Ala Pro Ile His Gly Val Gly Met Gln Gly
225 230 235 240
His Met Ser Leu Ser His Pro Asp Phe Ala Glu Phe Glu Lys Ser Ile
245 250 255
Ile Glu Tyr Ala Lys Leu Gly Val Lys Val His Val Thr Glu Leu Asp
260 265 270
Ile Asp Val Leu Pro Leu Ala Trp Asn Leu Ser Ala Glu Ile Ser Asn
275 280 285
Arg Phe Glu Tyr Arg Pro Glu Met Asp Pro Tyr Arg Glu Gly Leu Pro
290 295 300
Ala Lys Val Glu Glu Glu Leu Ala Ala Arg Tyr Glu Ala Leu Phe Lys
305 310 315 320
Ile Leu Leu Arg His Arg Asp Lys Ile Glu Arg Val Thr Thr Trp Gly
325 330 335
Thr Asn Asp Ser Glu Thr Trp Leu Asn Gly Phe Pro Ile Pro Gly Arg
340 345 350
Met Asn Tyr Pro Met Leu Phe Asp Arg Asn Asn Gln Pro Lys Leu Ala
355 360 365
Tyr His Arg Leu Leu Ala Leu Lys Gln Lys Lys Ser Gln
370 375 380

<210> 327
<211> 1500
<212> DNA
<213> Unknown

<220>
<223> Obtained from an environmental sample.

<400> 327
atgaaacggt cagtctctat ctttatcgca tgttttagtaa tgacagtatt aacaattagc 60
gggtgtcgcgg caccagaagc atctgcagca ggggcgaaaa cgccctgtagc ccttaatggc 120
cagcttagca ttaaagggtac tcagctagtc aatcaaaacg gaaaatcggg gcagctgaag 180
gggatcagct cacacgggttt gcagtgggttc ggcgattatg tcaataaaga ctctttaaaa 240
tggctaagag acgattgggg aattaccgtc ttccgagcgg caatgtacac ggctgaaggc 300
ggttatatag agaatccgtc tgtgaaaaat aaagtcaaaag aagctgttga agcggcaaaa 360
gagctcggga tatatgtcat cattgactgg catattttta atgacggcaa tccaaatcaa 420
aataaagaga aggcgaagga attctttaag gaaatgtcga gcctttacgg aagcacacca 480
aacgttattt atgaaattgc taatgaaccg aacgggtgatg taaattggaa gcgcgatatc 540

aaaccgtatg	cggaggaagt	gatttccggt	atccgtaaaa	atgaccgga	taacatcatt	600
attaccggaa	ctggcacttg	gagtcaggat	gtcaatgatg	ctgctgatga	tcagcttaag	660
gatgcaaacg	tcatgtacgc	gcttcatttt	tatgcaggta	cacacggcca	gtatttaagg	720
gataaagccg	attatgcgct	cagcaaagga	gcgccgattt	ttgtaacgga	atgggggacg	780
agtgacgctt	ccggaaatgg	cgggggtcttc	cttgaccagt	cgaggggaatg	gctgaattat	840
ctcgacaaca	agaaaatcag	ctgggtaaac	tgggaaccttt	ctgataagca	ggaatcttcc	900
tcagcttttaa	agccgggggc	atctaaaaca	ggcggctggc	cgttatcaga	tttatccgct	960
tcagggacat	ttgtaaggga	aaagatccgt	ggctcccaac	attcgactga	agacagatct	1020
gagacaccaa	agcaagataa	acccgtacag	gaaaacagcc	tatctgtgca	atacagaaca	1080
ggggatggaa	gtgtgaacag	caaccaaadc	cgctctcaga	tccatgtgaa	aaacaacagc	1140
aagaccaccg	ttattttaaa	aaatgtaact	gtccgctact	ggtataaacac	gaaaaacaaa	1200
ggccaaaact	tcgactgtga	ctacgcgaag	atcggtatgca	gcaatgtgac	gcacaagttt	1260
gtgacattac	aaaaacctgt	aaaaggtgca	gatgcctatc	tgggaacttg	gtttaaaaaac	1320
gggacactgt	caccgggagc	aaacactgga	gaaatccaaa	ttcgtcttca	caatgaggat	1380
tggggcaatt	attcacaat	cggggattat	tctttttctc	agtcaaatat	gtttaaagat	1440
acaaaaaaaa	tcacattata	taataacgga	aaactaat	ggggaactga	acccaaatag	1500

<210> 328

<211> 499

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(29)

<400> 328

Met	Lys	Arg	Ser	Val	Ser	Ile	Phe	Ile	Ala	Cys	Leu	Val	Met	Thr	Val
1				5					10					15	
Leu	Thr	Ile	Ser	Gly	Val	Ala	Ala	Pro	Glu	Ala	Ser	Ala	Ala	Gly	Ala
			20					25					30		
Lys	Thr	Pro	Val	Ala	Leu	Asn	Gly	Gln	Leu	Ser	Ile	Lys	Gly	Thr	Gln
		35					40					45			
Leu	Val	Asn	Gln	Asn	Gly	Lys	Ser	Val	Gln	Leu	Lys	Gly	Ile	Ser	Ser
	50				55					60					
His	Gly	Leu	Gln	Trp	Phe	Gly	Asp	Tyr	Val	Asn	Lys	Asp	Ser	Leu	Lys
65					70				75					80	
Trp	Leu	Arg	Asp	Asp	Trp	Gly	Ile	Thr	Val	Phe	Arg	Ala	Ala	Met	Tyr
			85					90						95	
Thr	Ala	Glu	Gly	Gly	Tyr	Ile	Glu	Asn	Pro	Ser	Val	Lys	Asn	Lys	Val
			100					105					110		
Lys	Glu	Ala	Val	Glu	Ala	Ala	Lys	Glu	Leu	Gly	Ile	Tyr	Val	Ile	Ile
	115						120					125			
Asp	Trp	His	Ile	Leu	Asn	Asp	Gly	Asn	Pro	Asn	Gln	Asn	Lys	Glu	Lys
	130					135					140				
Ala	Lys	Glu	Phe	Phe	Lys	Glu	Met	Ser	Ser	Leu	Tyr	Gly	Ser	Thr	Pro
145					150					155				160	
Asn	Val	Ile	Tyr	Glu	Ile	Ala	Asn	Glu	Pro	Asn	Gly	Asp	Val	Asn	Trp
			165					170						175	
Lys	Arg	Asp	Ile	Lys	Pro	Tyr	Ala	Glu	Val	Ile	Ser	Val	Ile	Arg	
			180					185					190		
Lys	Asn	Asp	Pro	Asp	Asn	Ile	Ile	Ile	Thr	Gly	Thr	Gly	Thr	Trp	Ser
	195					200						205			
Gln	Asp	Val	Asn	Asp	Ala	Ala	Asp	Asp	Gln	Leu	Lys	Asp	Ala	Asn	Val
	210				215						220				
Met	Tyr	Ala	Leu	His	Phe	Tyr	Ala	Gly	Thr	His	Gly	Gln	Tyr	Leu	Arg
225					230				235					240	
Asp	Lys	Ala	Asp	Tyr	Ala	Leu	Ser	Lys	Gly	Ala	Pro	Ile	Phe	Val	Thr
			245						250				255		
Glu	Trp	Gly	Thr	Ser	Asp	Ala	Ser	Gly	Asn	Gly	Gly	Val	Phe	Leu	Asp
			260					265					270		
Gln	Ser	Arg	Glu	Trp	Leu	Asn	Tyr	Leu	Asp	Asn	Lys	Lys	Ile	Ser	Trp
	275					280						285			
Val	Asn	Trp	Asn	Leu	Ser	Asp	Lys	Gln	Glu	Ser	Ser	Ser	Ala	Leu	Lys
	290					295					300				
Pro	Gly	Ala	Ser	Lys	Thr	Gly	Gly	Trp	Pro	Leu	Ser	Asp	Leu	Ser	Ala
305					310					315					320

Ser Gly Thr Phe Val Arg Glu Lys Ile Arg Gly Ser Gln His Ser Thr
 325 330 335
 Glu Asp Arg Ser Glu Thr Pro Lys Gln Asp Lys Pro Val Gln Glu Asn
 340 345 350
 Ser Leu Ser Val Gln Tyr Arg Thr Gly Asp Gly Ser Val Asn Ser Asn
 355 360 365
 Gln Ile Arg Pro Gln Ile His Val Lys Asn Asn Ser Lys Thr Thr Val
 370 375 380
 Asn Leu Lys Asn Val Thr Val Arg Tyr Trp Tyr Asn Thr Lys Asn Lys
 385 390 395 400
 Gly Gln Asn Phe Asp Cys Asp Tyr Ala Lys Ile Gly Cys Ser Asn Val
 405 410 415
 Thr His Lys Phe Val Thr Leu Gln Lys Pro Val Lys Gly Ala Asp Ala
 420 425 430
 Tyr Leu Glu Leu Gly Phe Lys Asn Gly Thr Leu Ser Pro Gly Ala Asn
 435 440 445
 Thr Gly Glu Ile Gln Ile Arg Leu His Asn Glu Asp Trp Gly Asn Tyr
 450 455 460
 Ser Gln Ile Gly Asp Tyr Ser Phe Ser Gln Ser Asn Thr Phe Lys Asp
 465 470 475 480
 Thr Lys Lys Ile Thr Leu Tyr Asn Asn Gly Lys Leu Ile Trp Gly Thr
 485 490 495
 Glu Pro Lys

<210> 329
 <211> 2268
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 329
 atgaggaacg tttaggaat aggaggcagt atgtacaaaa aggcttttct tgtactggca 60
 ttgtttttgc ttttggcggc ggtggcgctc ccgtctgtgg gggctgcgcc gcagggggccg 120
 cgcctgcgcg atgtggcggg cgacatttta gtgggttacg cctccagaaa cgatttctgg 180
 aacatgtctg actcagccca atacacagaa gttgcccgcg ctgagttcaa cttcatgacg 240
 cccgaaaacg ccatgaagt ggacgccatt catcccgcgc aaaactcata cagttttgcc 300
 caggccgacc ggcacgtgca gtttggccag gccacaaca tggccgtgca tggacatgcc 360
 ctctgtgtggc acagccaaaa tccaggctgg ctgaccaatg gcaactggtc ccgcagccaa 420
 ttgatcaaca tcatgaacga ccacattgac acggtcgccg gccgttatgc aggtgaggtg 480
 ctggtgtggg acgtggtcaa tcaggcggtt aatgaggatg gaacttatcg cagcaccatc 540
 tggtaacaa ggtatggaca ggaatatatc gacctggcct ttaccgcgc ccgcgccgcc 600
 gatcctcatg ccaaactcat ttacaacgat tacaacattg gctggttaaa cagtaagtcg 660
 aatggcgctc acaacatggc cgccgatatg gtcaggcgcg gtgtgcccac cgacggcggt 720
 ggtttccaga tgcacctgga acggggcggc gtcagcgcca gcagtctggc gagcaacatg 780
 cagcgtttgc ccgatttggg attggaagt tacatcacg aattggacgt gcgcattccc 840
 caaaacccaa cccagcagga tttgcaggct caggcgccag tttaacaaac ggtgacgaat 900
 cgctgtttgg cgagcctgc ctgcaaggcg ttgcaggctt ggggcatccc cgacaaatat 960
 tcctgggtac cggacgtatt ccccggcacg ggcgcgccct tgttgtttaa cgacaatat 1020
 gaggccaaac ccgctatta tgccgtccag gcagagttga tggccgcgaa tccgcagccc 1080
 acaaacacac cgggaacgcc cgctcatacc ctttcggcca cgtctacgtc tgcggccact 1140
 gctacgcccc cggcaacggc cacggcgacc gccaccaccc cctccggcgg cggcgtttgc 1200
 gccgttgatt acgtcattgc caaccagtgg ggcaatggct ttcaggccaa cgtcaccatc 1260
 accaatcaca gcgcccgcgc ggtgaacggc tataccctgg cctggacca cgcgccgggg 1320
 cagattgtca ccagcgctg gaacgtaac atcgccaaa gcggcagcgc cgtcagcgcc 1380
 agcaaccggg ccggttattg gaacggtgtg atcggagcca acggcgccaa gatttctttt 1440
 gggtttccagg gatctctggc gggcggcagc gcggtcgcg ccacttattt tgcctgaac 1500
 ggcgtgctt gtaacggggc cgtccttccg cttactgcca cttcacgcc ttcaccgacg 1560
 gctaccatgt gtcccaggc aacgcctgaa ctgcttgcg tgcagccggt gacttcaccc 1620
 actaccaaac tgtctcaaac gctgtgtgtg gctttaggca acggcgaatg ggtgcgcgct 1680
 gccggaccgg caggcgttgt caccgtcact gcgcccggacc cggatgggta tttccgcctg 1740
 acgataccgc tggcagccaa taccagcaac gccattctgg tagaaggcg ggtgcgggtt 1800
 atcacccatt caaatggctg cacctatggc ggttatacct tgagcagaac cgtaacgatt 1860
 gtgcaagcca gcagcccagt cacctaacg ccagatgcca cactttcccc caccgccacg 1920
 gcaacgccta cggtaaccgc cacgtcgccg tcaggcgccg gcaccgtcgc ctacgccatc 1980
 accaagcact ggggcagcgg ttaccgccg aacgttacc tcaccaatac tggcggaagc 2040
 gccctcaacg gctggaccct ggcctatgcc tttcccgcca atcaaaccat cagcaacgcc 2100

```

tggaacggaa cggccgttca gtccggcagc agcgctcagcg tcaccaacgc cggttggaat 2160
ggcagcctgc cgcccaacgt ctccgccagc ttgggttcc aggcgagcta cagcggcaat 2220
aacagcgtcc ctgccagctt tacgctgaac ggcgcgcttt gccattga 2268

```

<210> 330
 <211> 755
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(35)

<400> 330
 Met Arg Asn Val Gln Glu Ile Gly Gly Ser Met Tyr Lys Lys Ala Phe
 1 5 10 15
 Leu Val Leu Ala Leu Phe Leu Leu Leu Ala Ala Val Ala Leu Pro Ser
 20 25 30
 Val Gly Ala Ala Pro Gln Gly Pro Arg Leu Arg Asp Val Ala Gly Asp
 35 40 45
 Ile Leu Val Gly Tyr Ala Ser Arg Asn Asp Phe Trp Asn Met Ser Asp
 50 55 60
 Ser Ala Gln Tyr Thr Glu Val Ala Arg Thr Glu Phe Asn Phe Met Thr
 65 70 75 80
 Pro Glu Asn Ala Met Lys Trp Asp Ala Ile His Pro Ala Gln Asn Ser
 85 90 95
 Tyr Ser Phe Ala Gln Ala Asp Arg His Val Gln Phe Ala Gln Ala Asn
 100 105 110
 Asn Met Ala Val His Gly His Ala Leu Val Trp His Ser Gln Asn Pro
 115 120 125
 Gly Trp Leu Thr Asn Gly Asn Trp Ser Arg Ser Gln Leu Ile Asn Ile
 130 135 140
 Met Asn Asp His Ile Asp Thr Val Ala Gly Arg Tyr Ala Gly Glu Val
 145 150 155 160
 Leu Val Trp Asp Val Val Asn Gln Ala Phe Asn Glu Asp Gly Thr Tyr
 165 170 175
 Arg Ser Thr Ile Trp Tyr Asn Gly Ile Gly Gln Glu Tyr Ile Asp Leu
 180 185 190
 Ala Phe Thr Arg Ala Arg Ala Ala Asp Pro His Ala Lys Leu Ile Tyr
 195 200 205
 Asn Asp Tyr Asn Ile Gly Trp Leu Asn Ser Lys Ser Asn Gly Val Tyr
 210 215 220
 Asn Met Ala Ala Asp Met Val Arg Arg Gly Val Pro Ile Asp Gly Val
 225 230 235 240
 Gly Phe Gln Met His Leu Glu Arg Gly Gly Val Ser Gly Ser Ser Leu
 245 250 255
 Ala Ser Asn Met Gln Arg Phe Ala Asp Leu Gly Leu Glu Val Tyr Ile
 260 265 270
 Thr Glu Leu Asp Val Arg Ile Pro Gln Asn Pro Thr Gln Gln Asp Leu
 275 280 285
 Gln Ala Gln Ala Ala Val Tyr Gln Thr Val Thr Asn Arg Cys Leu Ala
 290 295 300
 Gln Pro Ala Cys Lys Ala Leu Gln Val Trp Gly Ile Pro Asp Lys Tyr
 305 310 315 320
 Ser Trp Val Pro Asp Val Phe Pro Gly Thr Gly Ala Pro Leu Leu Phe
 325 330 335
 Asn Asp Asn Tyr Glu Ala Lys Pro Ala Tyr Tyr Ala Val Gln Ala Glu
 340 345 350
 Leu Met Ala Ala Asn Pro Gln Pro Thr Asn Thr Pro Gly Thr Pro Ala
 355 360 365
 His Thr Pro Ser Ala Thr Ser Thr Ser Ala Ala Thr Ala Thr Pro Pro
 370 375 380
 Ala Thr Ala Thr Ala Thr Ala Thr Thr Pro Ser Gly Gly Gly Val Cys
 385 390 395 400
 Ala Val Asp Tyr Val Ile Ala Asn Gln Trp Gly Asn Gly Phe Gln Ala
 405 410 415
 Asn Val Thr Ile Thr Asn His Ser Ala Ala Pro Val Asn Gly Tyr Thr

115 120 125
 Gly Tyr Gln Trp Ile Gly Asn Thr Thr Ile Ser Thr Glu Ser Trp Thr
 130 135 140
 Thr Ile Arg Gly Thr Trp Leu Pro Arg Ala Asp Ala Asn Ala Ser Glu
 145 150 155 160
 Leu Tyr Val Tyr Pro Glu Val Thr Pro Val Ala Gly Phe Asp Tyr Leu
 165 170 175
 Leu Asp Asp Leu Leu Ile Glu Arg Ala Ala Pro Val Asp Gly Gly Ala
 180 185 190
 Pro Gly Thr Val Val Tyr Thr Ala Gly Phe Glu Thr Asp Leu Asp Gly
 195 200 205
 Trp Glu Ala Arg Ala Asp Gly Val Gly Val Gly Gln Leu Asp Arg Thr
 210 215 220
 Asp Ala Glu Ser Ala Glu Gly Asp Trp Ser Ala Ile Val Thr Asp Arg
 225 230 235 240
 Thr Ser His Gly His Gly Leu Arg Leu Asp Val Thr Asp Ile Met Asp
 245 250 255
 Ala Gly Val Thr Tyr Glu Ile Ser Ala Gln Val Lys Phe Ala Gly Thr
 260 265 270
 Gly Gly Pro Gly Asn Ile Trp Leu Ser Gln Glu Leu Val Val Asp Gly
 275 280 285
 Gly Ser Thr Tyr Gly Thr Val Leu Gln Val Pro Gly Val Thr Ser Thr
 290 295 300
 Ala Trp Thr Gln Ile Thr Thr Asn Tyr Val Thr Pro Thr Ala Asp Gln
 305 310 315 320
 Leu Phe Leu Tyr Phe Glu Thr Asn Trp Pro Asp Gly Ile Glu Asp Asp
 325 330 335
 Phe Leu Leu Asp Val Arg Ile Arg Val Ala Pro Arg Ala Ile Ile
 340 345 350
 Gln Glu Asp Leu Thr Pro Leu Met Asp Thr Leu Asp Val Pro Met Gly
 355 360 365
 Val Ala Ile Asp Gln Arg Glu Thr Ser Gly Ser Leu Ala Asp Leu Leu
 370 375 380
 Leu Leu His Phe Asp Gln Val Thr Ala Glu Asn His Met Lys Pro Glu
 385 390 395 400
 Ala Trp Tyr Asp Ala Ala Gly Asn Phe Arg Ile His Pro Gln Ala Arg
 405 410 415
 Ala Ile Met Asp Phe Ala Ala Glu Asn Asp Leu Arg Val Phe Gly His
 420 425 430
 Val Leu Val Trp His Gly Gln Thr Pro Asp Phe Phe Phe Thr His Ala
 435 440 445
 Asp Gly Thr Pro Leu Thr Ser Ser Glu Ala Asp Gln Ala Ile Leu Arg
 450 455 460
 Asp Arg Met Arg Thr His Ile Phe Asn Val Ala Glu Ala Leu Ser Glu
 465 470 475 480
 Trp Gly Glu Tyr Gly Asp Asn Pro Leu Val Ala Trp Asp Val Val
 485 490 495
 Asn Glu Val Val Ser Asp Ser Gly Glu His Ser Asp Gly Leu Arg Arg
 500 505 510
 Ser Arg Trp Tyr Asp Val Leu Gly Glu Phe Ile Asp Leu Ala Phe
 515 520 525
 Ile Tyr Ala Asn Gln Ala Phe Asn Gly Glu Phe Ala Ala Asp Asp Ala
 530 535 540
 Asn His Pro Val Thr Leu Phe Ile Asn Asp Tyr Asn Thr Glu Gln Ser
 545 550 555 560
 Gly Lys Gln Asn Arg Tyr Ala Ala Leu Ile Asp Arg Leu Ile Glu Arg
 565 570 575
 Glu Val Pro Ile Asp Ala Val Gly His Gln Phe His Val Ser Leu Ala
 580 585 590
 Met Pro Ile Ala Asn Leu Arg Gly Ala Leu Glu Arg Phe Gln Asp Thr
 595 600 605
 Gly Leu Ile Gln Gly Val Thr Glu Leu Asp Val Thr Val Gly Asn Asn
 610 615 620
 Pro Thr Glu Ala Leu Leu Val Glu Gln Gly Tyr Tyr Arg Asp Ala
 625 630 635 640
 Phe Arg Leu Phe Arg Glu Phe Thr Glu Asp Leu Tyr Ser Val Thr Val
 645 650 655
 Trp Gly Leu Thr Asp Asp Arg Ser Trp Arg Ser Ala Gln Ala Pro Leu
 660 665 670

Leu Phe Asp Ala Gly Leu Gln Ala Lys Pro Ala Tyr Tyr Gly Ala Ile
 675 680 685
 Asp Ala Asp Leu Asp Ala Arg Val Arg Ala Ala Tyr Val Phe Ala Glu
 690 695 700
 Asp Ile Ala Leu Asp Glu Ala Ala Leu Thr Ser Pro Thr Trp Asp Arg
 705 710 715 720
 Leu Pro Leu His Gln Ile Asp Gly Ala Gly Glu Phe Gln Leu Arg Trp
 725 730 735
 Ala Ala Asp His Leu Thr Val Phe Val His Val Thr Asp Gly Asp Glu
 740 745 750
 Val Glu Ile Val Leu Gly Asp Glu Thr Tyr Thr Val Ser Ser Asp Gly
 755 760 765
 Glu Gly Asp Leu Asp Ala Val Thr Ala Ala Gly Glu Asn Gly Ser Trp
 770 775 780
 Thr Ala Val Val Arg Val Pro Leu Thr Ala Glu Gln Gly Asp Thr Ala
 785 790 795 800
 Gln Phe Asp Leu Arg Ile Ile Asp Gly Ala Thr Thr Ser Gly Trp Asn
 805 810 815
 Val Glu Gly Val Leu Gly Thr Leu Thr Leu Val Glu Glu Leu Ser Phe
 820 825 830
 Val Glu Val Val Glu Ala Ala Asp Arg Pro Thr Ile Asp Gly Glu Ile
 835 840 845
 Asp Ala Val Trp Glu Asp Ala Asn Val Val Thr Thr Asp Val Arg Ile
 850 855 860
 Glu Gly Ala Ala Asp Gly Ala Lys Ala Glu Ile Arg Thr Leu Trp Asp
 865 870 875 880
 Asn Asn Thr Leu Phe Val Leu Ala Glu Ile Ala Asp Pro Val Ile Asp
 885 890 895
 Val Thr Ala Ser Ser Pro Trp Glu Gln Asp Ser Leu
 900 905

<210> 361
 <211> 5040
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 361
 atggcaagaa gtaagcgagt attagcatgg attatgtcta gtgtgcttct gatatccatg 60
 gcgatgccat ccttcgcatc aggtgattca agccaagtgc caagggttat atttgaaaca 120
 ggttttgaaa cgggattaga tggcttcaaa ggacggggta gtgccacctt aactcgaacg 180
 actgatgaaa cgcaagcagg cgactattcg gttcttgatga gcaatcggct tgagcactgg 240
 aatggggcat cattgccact tacaggcttc gttctaccag gtaatacata tgaatttggt 300
 ggttacataa aagcaaaagc agatgtagca gacaattatg tcatgagtgg tgagtacaat 360
 gaggggattt ctggaaatca atatccatgg atatctaatac gtttgtaaac ggttcaagat 420
 ggctttgttg agtttagagg tgaactaacc atactagagg atatgacgtc ctttaattcta 480
 aactttgaac atcaaaatgc tgaagtggaa ttttatitag attctgttca ggttatttcta 540
 atcgaagaag gtcaagtcaa tgacttacca atgaatgtaa gaagagcgcc acttacactt 600
 gctgaaactc ctttacatga gatttgggca gatcacttta ctattggcaa tattttatcg 660
 ccagggtttc gcacagatat acgtgggtgag gtattagccc atcatttttaa tgtgatcaca 720
 gctgaaaata ttatgaagcc agatcatttg caaagggaac aagggtatttt tacttttagt 780
 gcttccaacg atatgatgga atttgccaga gcaaataatc aagaagtcac tggacatact 840
 ttgggtgtggc atttcacatc cttcccatgg tttgaagcct taaatccaac acgtgatgaa 900
 gctatagcca ttatgcatac ccatattgaa actgttatgg gacatttttaa tgaaaactac 960
 ccagggtgtca ttacaggatg ggatgttttg aatgaagcca ttcaaccaag acagggtcaa 1020
 gatcctgaaa attggcgctt gcatttaagg gataccaaat gggtacgtgc cattggtgat 1080
 gatttatattg ccacgcgttt taacaaagcc catgaaatgg atccagatgc tattctttat 1140
 tataatgatt ataattgataa tgactatttt aaagcaacca ttataaaaagc catgggtgcag 1200
 gagttgcgta atgaaggcgt gccatttcat tgcaagggtca ttataattta 1260
 cagacaccat taaactctat tagaaccagc gttgagcgtt ttagtgaaat tactggtcat 1320
 gaagatctac cacctatttg cattagtttc acagaaattg atgtaacggc accagggttt 1380
 gaaagtgcag cccgtttacc tgaagaggta gaaattcgcc aagctcagtt ttatgctcaa 1440
 ttaatgcaga ttttaagaga caacagcgat cgtatttcac gtgttacttt ctgggggtatg 1500
 tctgatcgtg aatcatggcg tgcagatcgt catcctaaca tggtagatcc tcagtatggt 1560
 ccaaagcatg tctttcatgc tatagccaat ccaggagcct tccttacggc ctaccattta 1620
 ccagagacgc cagatgctca aacagcctat gcatctcaag gtcaaccagt tgtggggcag 1680
 ttaattttgg atgctgatca aaattcagaa gtaataccag tggctaatac aatgacagcc 1740

```

ctgtcgcgtgc gcgctccggct cgtgggcgac ggcacggcga cgacgccggc ccggtggatc 360
ggccgcgacc ccggagccga gaacggctac cagtggatcg gtaacacgac gatctcgacc 420
gagagctgga cgaccatccg ggaacgtgg ctccctcggg cggacgcgaa cgcctcggag 480
ctctatgtct accccgaggt cacaccggtg gccggcttcg actacctcct cgatgacctg 540
ctcatcgagc gtgctgcccc tgtcgacggc ggcgccccgg gcaccgtcgt ctacaccgct 600
ggattcgaga cggacctgga cggctgggag gcacgcgccg acggcgtcgg tgtcggccag 660
ctcgaccgga ccgacgcgga gtcggccgag ggcgactggt ccgcatcgt gaccgaccgc 720
acctcgcacg gccacggcct gcgcctggac gtcacggaca tcatggacgc gggcgtcacg 780
tacgagatca gcgccaggt gaagttcgcc gggacccggtg gtccgggcaa catctggctg 840
agccaggagc tggctcgtgga cgggggtagc acctacggca ccgtcctcca ggctcctggc 900
gtcacctcga cagcctggac gcagatcacc acgaactacg tcacgccgac ggccgaccag 960
ctgttctctt acttcgagac gaactggccg gacggcatcg aggacgactt cctcctcgac 1020
gacgtccgca tccgtgtcgc ccctcgggcg atcatccagg aggacctcac tccgctgatg 1080
gacacgctgg acgtgcccac ggggtgtcgc atcgaccagc gtgagacctc cggcagcctc 1140
gcggaccctc tgctgtgca cttcgaccag gtcacggccg agaaccacat gaagccggag 1200
gcgtggtacg accgcggcgg caacttcgc atccacccgc aggccgcgc catcatggag 1260
ttcgcggcgg agaacgacct gcgcgtcttc ggtcacgtcc tgggtgtggca cggccagacc 1320
ccggacttct tcttcacgca cgcggacggc accccgctga cctcgagcga ggccgaccag 1380
cgatcctgc gcgaccgcat gcgcacgcac atcttcaacg tcgccgaggc cctctccgag 1440
tggggcgagt tggcgccgga caaccgctc gtggcctggg acgtcgtcaa cgaggtcgtc 1500
tccgacagcg gcgagcacag cgacggcctg cgccgtagcc gctggtacga cgtgctgggc 1560
gaggagttca tcgacctggc gttcatctac gcgaaccagg cgttcaacgg tgagttcgcc 1620
gctgacgacg ccaaccaccc ggtcacgctc ttcataacg actacaacac cgagcagtc 1680
ggcaagcaga accggtacgc cgcgctcatc gaccgcctca tcgagcgcga ggtcccgatc 1740
gacgccgtgg ggcaccagtt ccacgtcagc ctggccatgc ccatcgcgaa cctgcgcggc 1800
gccctcgagc gcttcagga caccgggctg atccagggcg tcaccgagct cgacgtcacc 1860
gtcggcaaca acccgaccga ggcgctgctc gtcgagcagg gctactacta ccgggacgcc 1920
ttccggctgt tccgtgagtt cacggaggac ctctactcgg tcaccgtgtg gggctcacc 1980
gacgaccgca cgtggcgag gctcaggcg ccgctgctgt tcgacgcggg cctgcaggcc 2040
aagccggcct actacggcg ccatcgacgc gacctggacg cacgcgtcgc tgcggcctac 2100
gtgttcgccg aggacatcg cctcgacgag gccgcgctga cgagccccac ctgggaccgt 2160
ctgccgctgc accagatcga cggggccggc gagttccagc tccgctgggc ggccgaccac 2220
ctcacggtgt tcgtccacgt caccgacggt gacgaggtcg agatcgtgct cggcgacgag 2280
acctacacg tctcgtcgga cggcgagggc gacctggacg cggtcaccgc ggccggggag 2340
aacggctcct ggaccgctgt ggtccgcgtg ccgctcacgg ccgagcaggg cgacaccgcc 2400
cagttcgacc tccggatcat cgacggcgcc accacctccg ggtggaacgt cgaaggtgtc 2460
ctgggcaccc tgacctggt cgaggagctg tccttcgtcg aggtcgtcga ggcggccgac 2520
cggccgacca tcgacggcga gatcgacgcc gtytgggagg acgccaacgt cgtcaccacg 2580
gacgtccgta tcgaggcg cgtgacggc gcgaaggccg agatccggac cctgtgggac 2640
aacaacacgc tgttcgtcct cgcggagatc gccgacccgg tgatcgacgt gacggcctcc 2700
agcccgtggg agcaggactc gctc

```

<210> 360
 <211> 908
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(31)

<400> 360
 Met Thr Arg Ser Val Arg Pro Arg Ala Trp Gly Ala Gly Leu Leu Ala
 1 5 10 15
 Leu Ala Met Val Ala Thr Val Ala Pro Thr Ala Thr Gly His Ser His
 20 25 30
 Asp Thr Ala Glu Pro Val Val Val Tyr Thr Asp Phe Glu Asn Asp
 35 40 45
 Ser Ile Glu Pro Trp Ala Gln Ser Gly Gly Pro Thr Leu Asn Ile Val
 50 55 60
 Glu Val Asp Gly Gly His Ala Leu Arg Val Gly Asn His Gln Asn Thr
 65 70 75 80
 Trp Asp Gly Ile Gln Thr Gln Pro Ala Thr Thr Arg Ile Glu Pro Gly
 85 90 95
 Val Glu His Thr Leu Ser Met Arg Val Arg Leu Val Gly Asp Gly Thr
 100 105 110
 Ala Thr Thr Pro Ala Arg Trp Ile Gly Arg Asp Pro Gly Ala Glu Asn

<210>	353
<211>	1983
<212>	DNA
<213>	Unknown

<220>
<223> obtained from an environmental sample.

<210>	354
<211>	660
<212>	PRT
<213>	Unknown

<220>
<223> obtained from an environmental sample.

<221> SIGNAL
<222> (1)...(18)

<400> 354
Met Val Cys Ser Ala Ala Leu Gly Ala Thr Ala Val Ser Ala Gln Thr
1 5 10 15
Leu Ser Asn Asn Ser Thr Gly Thr Asn Gly Phe Tyr Tyr Thr Phe
20 25 30

580
 Ser Ser Thr Pro Val Val Val Ser Ser Ser Arg Ser Ser Ser Ser Val
 595 600 605
 Ala Ala Gly Gly Ala Cys Gln Cys Asn Trp Trp Gly Thr Arg Tyr Pro
 610 615 620
 Ile Cys Thr Thr Thr Ala Ser Gly Trp Gly Trp Glu Asn Asn Arg Ser
 625 630 635 640
 Cys Ile Thr Thr Ser Thr Cys Asn Ser Gln Gly Pro Gly Gly Gly Gly
 645 650 655
 Val Val Cys Asn
 660

<210> 355
 <211> 1125
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 355
 atgaaacgaa ccattcttctt aagactctta gccggcgccc tcctctccgc cgcggccctc 60
 gcggccgggg ggtgccggcc ctctagtccc ccgaaggctc agatcgaggc caatatcccc 120
 tccctcaaag aggtctgcgc ttcttatttc gagatcggcg cggccgtcga gccgtatcag 180
 ttatctcttc caccacacga tgcccttctg cggaaacatt ttaactgcct cgtggcggag 240
 aacgtcatga agcccgctc catccagcct tcggaggggt atttcaactg gaccgaagca 300
 gacaagatcg tgaactacgc caaagcccac gggatgaagc tccgcttcca taccctcgtc 360
 tggcataatc aggtcccga ttggttcttc gcgggtaacg acaaaaccct ctttttgcag 420
 cgcttgagaga atcatatccg gactatcatt aaaagatatg gcgataaggt cgactattgg 480
 gacgtggtaa atgaggctat agaccgcagc caaccggatg gcatgaggag gagcaaattg 540
 taccagatca ccgggaagga ctacatcaag accgccttcc ggggtggcaga cgacgagctc 600
 aggaagaatg ggtggaggaa agaaggctcgt cagctctata tcaacgacta caacacccat 660
 gatccgacga agagagagta catctggcgc ttgatcgatg agcttcaaac ggaagggatt 720
 cccgtcgacg gagtaggcca ccagacgcat atcaatatcg aatggccgcc cgtaaaccag 780
 atcgtggact cgatccgctt cttcggggaa aaaggcctcg ataaccaggt gaccgagctg 840
 gatgtgagca tatatacggg tagatccagt tcctacggga gttaccaagc gatcccgag 900
 gaagtcttca tcaagcaggg taatcgctac aaggaactct ttgaagggct aaaaagtgtg 960
 aaaaactacc tcagcaacgt caccttctgg ggcattggcg acgatcatac ctggctgaac 1020
 cattggccca tcgaacggcc cgatgctcct cttcttctcg atatctatct caaggccaag 1080
 ccggcgatatt gggggatcgt ggatgctttg aagctttcgc ggtga 1125

<210> 356
 <211> 374
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(21)

<400> 356
 Met Lys Arg Thr Ile Phe Leu Arg Leu Leu Ala Gly Ala Leu Leu Ser
 1 5 10 15
 Ala Ala Ala Leu Ala Ala Gly Gly Cys Arg Pro Ser Ser Pro Pro Lys
 20 25 30
 Val Glu Ile Glu Ala Asn Ile Pro Ser Leu Lys Glu Val Cys Ala Ser
 35 40 45
 Tyr Phe Glu Ile Gly Ala Ala Val Glu Pro Tyr Gln Leu Ser Ser Pro
 50 55 60
 Pro His Asp Ala Leu Leu Arg Lys His Phe Asn Cys Leu Val Ala Glu
 65 70 75 80
 Asn Val Met Lys Pro Ala Ser Ile Gln Pro Ser Glu Gly Tyr Phe Asn
 85 90 95
 Trp Thr Glu Ala Asp Lys Ile Val Asn Tyr Ala Lys Ala His Gly Met
 100 105 110
 Lys Leu Arg Phe His Thr Leu Val Trp His Asn Gln Val Pro Asp Trp
 115 120 125

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(25)

<400> 358

```

Met Asn Asn Phe Arg Asn Thr Phe Leu Ile Val Val Val Leu Ala Val
 1      5      10      15
Val Val Gly Val Leu Pro Ala Cys Glu Ala Gly Pro Pro Glu Asn Thr
 20      25      30
Ser Ser Ser Leu Gln Glu Ala Tyr Ala Asp Val Phe Leu Ile Gly Thr
 35      40      45
Ala Leu Asn Leu Ala Gln Ile Asp Gly Arg Asp Glu Gln Gly Val Arg
 50      55      60
Leu Val Glu Arg His Phe Asn Ala Ile Thr Pro Glu Asn Ile Thr Lys
 65      70      75      80
Trp Gly Pro Ile His Pro Ala Pro Gly Glu Tyr Asn Phe Gly Pro Ala
 85      90      95
Asp Arg Phe Val Glu Phe Gly Glu Ala His Asp Met Phe Met Ile Gly
 100     105     110
His Thr Leu Val Trp His Ser Gln Thr Pro Gly Trp Val Phe Glu Asp
 115     120     125
Glu Ala Gly Asn Pro Leu Gly Arg Asp Glu Leu Ile Glu Arg Met Arg
 130     135     140
Asp His Ile His Thr Val Val Gly Arg Tyr Arg Gly Arg Ile His Ala
 145     150     155     160
Trp Asp Val Val Asn Glu Ala Leu Asn Glu Asp Gly Thr Leu Arg Glu
 165     170     175
Ser Pro Trp Tyr Arg Ile Ile Gly Glu Asp Tyr Leu Leu Lys Ala Phe
 180     185     190
Glu Phe Ala His Glu Ala Asp Pro Asp Ala Glu Leu Tyr Tyr Asn Asp
 195     200     205
Tyr Ser Leu Glu Asn Pro Ala Lys Arg Ala Gly Ala Val Arg Leu Val
 210     215     220
Arg Tyr Leu Gln Glu Asn Gly Ala Pro Ile His Gly Ile Gly Thr Gln
 225     230     235     240
Gly His Tyr Ser Leu Asp Trp Pro Ser Leu Asp Glu Ile Glu Arg Thr
 245     250     255
Ile Thr Asp Phe Ala Ala Leu Asp Val Asp Val Met Val Thr Glu Leu
 260     265     270
Glu Ile Asp Val Leu Pro Ser Ala Phe Glu Tyr Gln Gly Ala Asp Ile
 275     280     285
Ala Met Arg Ala Glu Leu Glu Glu Arg Leu Asn Pro Tyr Pro Asp Glu
 290     295     300
Leu Pro Ala Glu Val Asp Glu Ala Leu Thr Gln Arg Tyr Arg Asp Ile
 305     310     315     320
Phe Glu Val Phe Leu Arg His Ser Asp Val Leu Thr Arg Val Thr Phe
 325     330     335
Trp Gly Val Thr Asp Gly Asp Ser Trp Lys Asn Asn Trp Pro Val Pro
 340     345     350
Gly Arg Thr Asn Tyr Pro Leu Leu Phe Asp Arg Glu Trp Gln Pro Lys
 355     360     365
Pro Ala Phe Tyr Ser Val Ile Glu Val Ala Asp Glu Met Leu Asn Glu
 370     375     380

```

<210> 359

<211> 2724

<212> DNA

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 359

```

atgacccggt ctgtacgacc aagagcatgg ggggccggcc tactggccct cgcgatggtg      60
gcgacagtcg cccccacggc caccggccac agccacgaca cggcagagcc cgtcgtcgtg      120
gtctacaccg acttcgagaa cgacagcatc gagccgtggg cgcagtccgg cggcccgacg      180
ctgaacatcg tcgaggtcga cggcgggcac gcgctgcgcg tcggcaacca ccagaacacc      240
tgggacggca tccagaccga gcccgccacc acgcggatcg agccgggtgt cgagcacacc      300

```

Phe Phe Ala Gly Asn Asp Lys Thr Leu Leu Leu Gln Arg Leu Glu Asn
 130 135 140
 His Ile Arg Thr Ile Ile Lys Arg Tyr Gly Asp Lys Val Asp Tyr Trp
 145 150 155 160
 Asp Val Val Asn Glu Ala Ile Asp Pro Ser Gln Pro Asp Gly Met Arg
 165 170 175
 Arg Ser Lys Trp Tyr Gln Ile Thr Gly Lys Asp Tyr Ile Lys Thr Ala
 180 185 190
 Phe Arg Val Ala Asp Asp Glu Leu Arg Lys Asn Gly Trp Arg Lys Glu
 195 200 205
 Gly Arg Gln Leu Tyr Ile Asn Asp Tyr Asn Thr His Asp Pro Thr Lys
 210 215 220
 Arg Glu Tyr Ile Trp Arg Leu Ile Asp Glu Leu Gln Thr Glu Gly Ile
 225 230 235 240
 Pro Val Asp Gly Val Gly His Gln Thr His Ile Asn Ile Glu Trp Pro
 245 250 255
 Pro Val Asn Gln Ile Val Asp Ser Ile Arg Phe Phe Gly Glu Lys Gly
 260 265 270
 Leu Asp Asn Gln Val Thr Glu Leu Asp Val Ser Ile Tyr Thr Asp Arg
 275 280 285
 Ser Ser Ser Tyr Gly Ser Tyr Gln Ala Ile Pro Gln Glu Val Phe Ile
 290 295 300
 Lys Gln Gly Asn Arg Tyr Lys Glu Leu Phe Glu Gly Leu Lys Ser Val
 305 310 315 320
 Lys Asn Tyr Leu Ser Asn Val Thr Phe Trp Gly Met Ala Asp Asp His
 325 330 335
 Thr Trp Leu Asn His Trp Pro Ile Glu Arg Pro Asp Ala Pro Leu Pro
 340 345 350
 Phe Asp Ile Tyr Leu Lys Ala Lys Pro Ala Tyr Trp Gly Ile Val Asp
 355 360 365
 Ala Leu Lys Leu Ser Arg
 370

<210> 357
 <211> 1155
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 357
 atgaataact tcagaaatac atttctaatac gtcgttgtag tggcggtcgt cgtcggcgtg 60
 ctgccggcct gcgaagccgg tccgccggaa aatacaagtt cgtccctgca ggaggcatat 120
 gcagatgtgt ttctgatcgg caccgcgctc aatctggcac agatcgacgg aagggatgaa 180
 caaggcgtag gtctggtgga gcggcatittt aatgcgatta caccagagaa cattacaaaa 240
 tggggaccga tacatccggc gccgggcgaa tataatttcg gaccggccga ccggtttgtt 300
 gaattcgggtg aagcccacga catgttcattg ataggccata cgcttgtagt gcacagccag 360
 acgcccggat ggggtattcga ggatgaagcc ggaaatccgc tcggccgcga cgagctcatc 420
 gaacgcattg cgcgatcatat ccataccgctc gtcggacggt accggggtag aatacacgca 480
 tgggacgtcg tcaacgaagc gttgaatgaa gacggaaccc tgcgggaatc cccctggtag 540
 cgtatcatcg gcgaggatta cctgttgaaa gcgttcgagt tcgcgcatga agcggacccg 600
 gatgccgagc tgtactataa cgattattct ctcgaaaatc ccgccaagcg ggcgggggag 660
 gtacgcctgg tccggtacat gcaggagaac ggggcgcgca tacacgggat cggtagccag 720
 ggacactact ctcttgactg gccatcgctc gacgagatcg aaagaaccat caccgatttc 780
 gccgcgttgg acgtggacgt catggttacc gaacttgaaa tcgacgtcct cccttcgag 840
 ttcgagtatc agggggccga tattgcatg cgggcggaac tcgaagagcg gttgaatccg 900
 tatcccgacg aactgccggc cgaggtcgat gaagcgctta cacagcggtg tcgggacatc 960
 ttcgaggtag ttctgcggca cagcgacgtt cttacgcgag taacgttctg ggggtgacc 1020
 gatggagatt cgtggaagaa taactggccg gtaccgggaa ggacgaatta tccgtgctg 1080
 ttcgaccgag aatggcagcc aaaaccagca ttttattccg tgatcgaagt tgcggatgag 1140
 atgctgaatg aataa 1155

<210> 358
 <211> 384
 <212> PRT
 <213> Unknown

<220>

Asp Phe Leu Met Trp Gly Phe Trp Glu Gly Ser His Trp Lys Pro Leu
 500 505 510
 Gly Ala Met Ile Arg Arg Asp Trp Ser Glu Lys Pro Met Tyr Arg Val
 515 520 525
 Trp Arg Glu Leu Ile Phe Glu Arg Trp Gln Thr Asp Glu Thr Gly Val
 530 535 540
 Thr Pro Glu His Gly Ala Ile Tyr Val Arg Gly Phe Lys Gly Asp Tyr
 545 550 555 560
 Glu Ile Thr Val Lys Ala Gly Gly Gln Glu Val Arg Val Pro Tyr Thr
 565 570 575
 Leu Lys Glu Asp Gly Gln Val Leu Trp Val Thr Val Gly Gly Thr Ser
 580 585 590
 Glu Glu Gln Ala Pro
 595

<210> 351
 <211> 1860
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 351
 atgcaccttc caaattaccg ctcgctcgca acggcgctca gtcgctacag ttgctcggcg 60
 ctgctggccg tcagtcttgt ggcctgcggg ggcaataatg accaagatcc gccgaccccg 120
 gagccaactc cgggtcccaac gccgactcca actccaaccc cgaccccgac tccggagcca 180
 accccgactc cgactccgga gccactcccg actccagagc caactccgac gccgaccccg 240
 gaaccaacac cgacgcggga gccgacgcca acaccggatc cgggggccga ctaccagccg 300
 cccagcaatg acattgccgt caatggcgac gtggaaagcg gtactaccaa ctgggggtgca 360
 cgcggttcgg catccattag ccgagtcact ttagagagct ttgaaggtga tgccagcttg 420
 agtgttaccg gccgagaaga cgactggcat ggcgccacct tctctgtagg ccatctgacc 480
 ccgggtaata gctatgaagt ggctgctgtg gtcaagttag cctcaggcga gcccaacaca 540
 gtggtcaaaa tcacgggtaa gcgcgagggc gagagcgcgga cttacgaaga gtacacggat 600
 gtcggtacgg cattggctac cgacggtagc tggaccgaaa ttaccggcac ttatatctct 660
 gatagcgcca gccatttga atattttatt gtggagaccc aagagggtgg accgaccgtt 720
 agcttctacg tggacgcgtt ttcagtggcc ggtgaggtgg aagatacgcc agcgccaacg 780
 ccgcccccaa ccgctccgcc accgagtggc tcaggcctag cggaactagt ggatttcccg 840
 gtgggcgttg ccgttgcggt agctagtttt gccataacg atttcctgag taacacgcaa 900
 caacaagata tcgtgcttaa caattttagt gaaattgttg cggagaatca gatgaagatg 960
 gaatttttca acgatgacta ctccaacccc agggcagatc aactggtcag ctggggccaat 1020
 gagcgaggta ttccgggtta cggccacgct ttggtctggc acgcgcaagc agcgctcatgg 1080
 gtcagtcctc cggtaagcaa ctttcgagag cgttatgtca accatgttcg tgggtgtggca 1140
 tcccggatg cggacacggt agtaagctgg gatgtcgtca atgaagcttt gaccgatgat 1200
 gatgtctccc cgggtggaag ctactaccgg caatctgagt tctaccgaca gttcaatggc 1260
 ccagagtcca ttgatctgc tttccgtgaa gctcagagag cagcccccaa tgcgctgctt 1320
 tactacaacg actacaatat tgagaacgga ctggacaaaa ccgatggttt gattcagcta 1380
 cttgagaggt tgagggataa tgacgtgcc attgatggcg tgggcttcca aatgcatggt 1440
 ttgttgatt ggcccgatat cagcactatt cgacgttcct gggagcgcgc attagcgggt 1500
 gaccccgatg accgtatgct tttaaaaata acggaactcg atgtgcgtat caacaatccc 1560
 tacgacgata atctcgaag aggcacgttt cactccagcc ggggtgactg cgatgacatt 1620
 tccgggggtc gcgaaggctt tgagcggcaa gccgctcggt accgggagat tattgaggcc 1680
 tactttgacg tcgtgccacc gcaccgccga ggtggcatca gtgtttgggg tattgccgac 1740
 cattatagtt ggtattatac ccatgaaggc tatgtcgatt ggcccctttt gtgggacagg 1800
 aacctacagc caaagcctgc ttacaacgct gtttatgaag ctctgcagca gggccaataa 1860

<210> 352
 <211> 619
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(73)

<400> 352
 Met His Leu Pro Asn Tyr Arg Ser Leu Ala Thr Ala Leu Ser Arg Tyr
 Page 266

1				5					10					15			
Ser	Cys	Ser	Ala	Leu	Leu	Ala	Val	Ser	Leu	Val	Ala	Cys	Gly	Gly	Asn		
			20					25					30				
Asn	Asp	Gln	Asp	Pro	Pro	Thr	Pro	Glu	Pro	Thr	Pro	Val	Pro	Thr	Pro		
		35					40					45					
Thr	Pro	Thr	Pro	Thr	Pro	Thr	Pro	Thr	Pro	Glu	Pro	Thr	Pro	Thr	Pro		
		50				55					60						
Thr	Pro	Glu	Pro	Thr	Pro	Thr	Pro	Glu	Pro	Thr	Pro	Thr	Pro	Thr	Pro		
65					70					75					80		
Glu	Pro	Thr	Pro	Thr	Pro	Glu	Pro	Thr	Pro	Thr	Pro	Asp	Pro	Gly	Ala		
				85				90						95			
Asp	Tyr	Gln	Pro	Pro	Ser	Asn	Asp	Ile	Ala	Val	Asn	Gly	Asp	Val	Glu		
			100					105					110				
Ser	Gly	Thr	Thr	Asn	Trp	Gly	Ala	Arg	Gly	Ser	Ala	Ser	Ile	Ser	Arg		
		115					120					125					
Val	Thr	Leu	Glu	Ser	Phe	Glu	Gly	Asp	Ala	Ser	Leu	Ser	Val	Thr	Gly		
		130				135					140						
Arg	Glu	Asp	Asp	Trp	His	Gly	Ala	Thr	Phe	Ser	Val	Gly	His	Leu	Thr		
145					150					155					160		
Pro	Gly	Asn	Ser	Tyr	Glu	Val	Ala	Ala	Trp	Val	Lys	Leu	Ala	Ser	Gly		
				165					170						175		
Glu	Pro	Asn	Thr	Val	Val	Lys	Ile	Thr	Gly	Lys	Arg	Glu	Gly	Glu	Ser		
			180					185					190				
Ala	Thr	Tyr	Glu	Glu	Tyr	Thr	Asp	Val	Gly	Thr	Ala	Leu	Ala	Thr	Asp		
		195					200					205					
Gly	Ser	Trp	Thr	Glu	Ile	Thr	Gly	Thr	Tyr	Ile	Pro	Asp	Ser	Ala	Ser		
		210				215					220						
Pro	Phe	Glu	Tyr	Phe	Ile	Val	Glu	Thr	Gln	Glu	Gly	Gly	Pro	Thr	Val		
225					230					235					240		
Ser	Phe	Tyr	Val	Asp	Ala	Phe	Ser	Val	Ala	Gly	Glu	Val	Glu	Asp	Thr		
				245					250					255			
Pro	Ala	Pro	Thr	Pro	Pro	Pro	Thr	Ala	Pro	Pro	Pro	Ser	Gly	Ser	Gly		
			260					265					270				
Leu	Ala	Glu	Leu	Val	Asp	Phe	Pro	Val	Gly	Val	Ala	Val	Ala	Val	Ala		
		275					280					285					
Ser	Phe	Ala	Asn	Asn	Asp	Phe	Leu	Ser	Asn	Thr	Gln	Gln	Gln	Asp	Ile		
		290				295					300						
Val	Leu	Asn	Asn	Phe	Ser	Glu	Ile	Val	Ala	Glu	Asn	Gln	Met	Lys	Met		
305					310					315					320		
Glu	Tyr	Phe	Asn	Asp	Asp	Tyr	Ser	Asn	Pro	Arg	Ala	Asp	Gln	Leu	Val		
				325					330					335			
Ser	Trp	Ala	Asn	Glu	Arg	Gly	Ile	Arg	Val	His	Gly	His	Ala	Leu	Val		
			340					345					350				
Trp	His	Ala	Gln	Ala	Ala	Ser	Trp	Val	Ser	Pro	Pro	Val	Ser	Asn	Phe		
		355				360						365					
Arg	Glu	Arg	Tyr	Val	Asn	His	Val	Arg	Gly	Val	Ala	Ser	Arg	Tyr	Ala		
		370				375					380						
Asp	Thr	Val	Val	Ser	Trp	Asp	Val	Val	Asn	Glu	Ala	Leu	Thr	Asp	Asp		
385					390					395					400		
Asp	Val	Ser	Pro	Gly	Gly	Ser	Tyr	Tyr	Arg	Gln	Ser	Glu	Phe	Tyr	Arg		
				405					410					415			
Gln	Phe	Asn	Gly	Pro	Glu	Phe	Ile	Asp	Ile	Ala	Phe	Arg	Glu	Ala	Arg		
			420					425					430				
Glu	Ala	Ala	Pro	Asn	Ala	Leu	Leu	Tyr	Tyr	Asn	Asp	Tyr	Asn	Ile	Glu		
		435				440						445					
Asn	Gly	Leu	Asp	Lys	Thr	Asp	Gly	Leu	Ile	Gln	Leu	Glu	Arg	Leu			
		450				455					460						
Arg	Asp	Asn	Asp	Val	Pro	Ile	Asp	Gly	Val	Gly	Phe	Gln	Met	His	Val		
465					470					475					480		
Leu	Leu	Asp	Trp	Pro	Asp	Ile	Ser	Thr	Ile	Arg	Arg	Ser	Trp	Glu	Arg		
				485					490					495			
Ala	Leu	Ala	Val	Asp	Pro	Asp	Asp	Arg	Met	Leu	Leu	Lys	Ile	Thr	Glu		
			500					505					510				
Leu	Asp	Val	Arg	Ile	Asn	Asn	Pro	Tyr	Asp	Asp	Asn	Leu	Glu	Arg	Gly		
		515					520					525					
Ile	Val	His	Ser	Ser	Arg	Gly	Asp	Cys	Asp	Asp	Ile	Ser	Gly	Val	Cys		
		530				535					540						
Glu	Gly	Phe	Glu	Arg	Gln	Ala	Ala	Arg	Tyr	Arg	Glu	Ile	Ile	Glu	Ala		
545					550					555					560		

Trp Lys Asp Ser Gly Ser Ala Thr Met Thr Leu Ala Ala Gly Gly Arg
 35 40 45
 Tyr Thr Ser Gln Trp Thr Asn Asn Thr Asn Asn Trp Val Gly Gly Lys
 50 55 60
 Gly Trp Asn Pro Gly Asn Ser Thr Arg Val Ile Ser Tyr Ser Gly Asn
 65 70 75
 Tyr Gly Val Ser Asn Ser Gln Asn Ser Tyr Leu Ala Leu Tyr Gly Trp
 85 90 95
 Thr Arg Ser Pro Leu Ile Glu Tyr Tyr Val Ile Glu Ser Tyr Gly Ser
 100 105 110
 Tyr Asn Pro Ala Ser Cys Ser Gly Gly Thr Asn Met Gly Ser Phe Gln
 115 120 125
 Ser Asp Gly Ala Thr Tyr Asp Val Arg Arg Cys Gln Arg Val Gln Gln
 130 135 140
 Pro Ser Ile Asp Gly Thr Gln Thr Phe Tyr Gln Tyr Phe Ser Val Arg
 145 150 155 160
 Asn Pro Lys Lys Gly Phe Gly Gln Ile Ser Gly Thr Ile Thr Phe Ala
 165 170 175
 Asn His Ala Ala Phe Trp Ala Ser Lys Gly Met Asn Leu Gly Ala His
 180 185 190
 Asn Tyr Gln Val Met Ala Thr Glu Gly Tyr Gln Ser Thr Gly Ser Ser
 195 200 205
 Asp Ile Thr Val Ser Glu Gly Pro Ile Asn Gly Gly Thr Ser Ser Thr
 210 215 220
 Pro Pro Val Thr Thr Ser Ser Ser Ala Ser Ser Val Ala Thr Gly Gly
 225 230 235 240
 Gly Asn Thr Gly Ser Gly Val Val Val Arg Ala Arg Gly Val Ala Gly
 245 250 255
 Gly Glu His Ile Asn Leu Arg Ile Gly Gly Asn Thr Val Ala Ser Trp
 260 265 270
 Asn Leu Thr Thr Ser Phe Gln Asp Leu Ser Tyr Ser Gly Thr Ala Ser
 275 280 285
 Gly Asp Ile Gln Val Gln Tyr Asp Asn Asp Gly Gly Ser Arg Asp Val
 290 295 300
 Val Val Asp Tyr Ile Arg Val Asn Gly Glu Thr Arg Gln Ala Glu Asp
 305 310 315 320
 Met Ser Tyr Asn Thr Ala Leu Tyr Ala Asn Gly Ser Cys Gly Gly Tyr
 325 330 335
 Gly Asn Ser Glu Leu Met His Cys Asn Gly Val Ile Gly Phe Gly Tyr
 340 345 350
 Thr Tyr Asp Cys Phe Ser Gly Asn Cys Ser Gly Gly Ser Thr Gly Gly
 355 360 365
 Gly Asn Thr Gly Thr Ser Ser Ser Ala Ala Ser Ala Gly Gly Gly Asn
 370 375 380
 Ser Asn Cys Ser Gly Tyr Val Gly Ile Thr Phe Asp Asp Gly Pro Thr
 385 390 395 400
 Ala Asn Thr Pro Thr Leu Val Asn Leu Leu Lys Gln Asn Asn Leu Thr
 405 410 415
 Pro Val Thr Trp Phe Asn Gln Gly Asn Asn Val Val Ala Asn Ala Asn
 420 425 430
 Tyr Met Ala Gln Gln Leu Ser Val Gly Glu Val His Asn His Ser Tyr
 435 440 445 450
 Ser His Pro Gln Met Gly Ser Met Thr Tyr Gln Gln Val Tyr Asp Glu
 455 460 465
 Leu Asn Arg Ala Asn Gln Ala Ile Gln Thr Ala Gly Ala Pro Lys Pro
 470 475 480
 Thr Leu Phe Arg Pro Pro Tyr Gly Thr Val Asn Ser Thr Ile Gln Gln
 485 490 495
 Ala Ala Gln Ala Leu Gly Leu Arg Val Ile Thr Trp Asp Val Asp Ser
 500 505 510
 Gln Asp Trp Asn Gly Ala Thr Ala Ser Ala Ile Ala Ser Ala Ala Asn
 515 520 525
 Arg Leu Thr Asn Gly Gln Val Ile Leu Met His Asp Gly Ser Tyr Thr
 530 535 540
 Asn Thr Asn Ala Ala Ile Ala Gln Ile Ala Ser Ser Leu Arg Ala Lys
 545 550 555 560
 Gly Leu Cys Pro Gly Arg Ile Asp Pro Ala Thr Gly Arg Ala Val Ala
 565 570 575
 Pro Ala Gly Gly Asn Thr Gly Gly Gly Thr Val Ser Ser Ser Thr Arg

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(20)

<400> 350

```

Met Pro Val Leu Phe Ala Leu Phe Leu Val Ala Ser Ser Cys Ala Ala
1      5      10      15
Gln Ser Leu Ala Gly Pro Val Ser Leu Leu Gly Gly Asp Ala Gly Ala
20      25      30
Ala Phe Arg Tyr Thr Gly Pro Ser Ala Gly Ala Ala Ser Gly Ser Ala
35      40      45
Glu Trp Val Ala Val Glu Asn Met Pro Phe Thr His Ala Trp Arg Leu
50      55      60
Arg Thr Asn Pro Leu Pro Glu Ser Gly Gly Asn Glu Trp Asp Leu Arg
65      70      75      80
Ile Arg Ala Arg Gly Ala Ala Ala Val Ser Ala Gly Asp Lys Ile Leu
85      90      95
Ala Glu Phe Trp Met Arg Cys Val Glu Pro Glu Asn Gly Asp Cys Ile
100      105      110
Leu Arg Leu Asn Val Glu Arg Asp Gly Ser Pro Trp Thr Lys Ser Ile
115      120      125
Ser Asn Pro Tyr Pro Val Gly Arg Glu Trp Arg Arg Phe Arg Val Leu
130      135      140
Phe Glu Met Arg Glu Ser Tyr Ala Ala Gly Gly Tyr Met Ile Asp Phe
145      150      155      160
Trp Met Gly Gln Gln Val Gln Thr Ala Glu Val Gly Gly Ile Ser Leu
165      170      175
Leu Asn Tyr Gly Pro Gln Ala Thr Ala Glu Gln Leu Gly Leu Asp Arg
180      185      190
Phe Tyr Glu Gly Ala Ala Ala Asp Ala Ala Trp Arg Gln Ala Ala Glu
195      200      205
Gln Arg Ile Glu Glu Ile Arg Lys Ala Gly Met Ile Ile Val Ala Val
210      215      220
Thr Pro Asp Gly Glu Pro Ile Glu Gly Ala Glu Ile Arg Ala Lys Leu
225      230      235      240
Lys Arg His Ala Phe Gly Trp Gly Thr Ala Val Ala Ala Ser Arg Leu
245      250      255
Leu Gly Thr Gly Thr Asp Ser Glu Arg Tyr Arg Asn Phe Ile Arg Glu
260      265      270
Asn Phe Asn Met Ala Val Leu Glu Asn Asp Leu Lys Trp Gly Pro Phe
275      280      285
Glu Glu Asn Arg Ala Arg Ala Met Asn Ala Leu Arg Trp Leu His Glu
290      295      300
Asn Gly Ile Thr Trp Ile Arg Gly His Asn Leu Val Trp Pro Gly Trp
305      310      315      320
Arg Trp Met Pro Ser Asp Val Arg Asn Leu Ala Asn Asn Pro Glu Ala
325      330      335
Leu Arg Gln Arg Ile Leu Asp Arg Ile Arg Asp Thr Ala Thr Ala Thr
340      345      350
Arg Gly Leu Val Val His Trp Asp Val Val Asn Glu Pro Val Ala Glu
355      360      365
Arg Asp Val Leu Asn Ile Leu Gly Asp Glu Val Met Ala Asp Trp Phe
370      375      380
Arg Ala Ala Lys Glu Cys Asp Pro Glu Ala Arg Met Phe Ile Asn Glu
385      390      395      400
Tyr Asp Ile Leu Ala Ala Asn Gly Ala Asn Leu Arg Lys Gln Asn Ala
405      410      415
Tyr Tyr Arg Met Ile Glu Met Leu Leu Lys Leu Glu Ala Pro Val Glu
420      425      430
Gly Ile Gly Phe Gln Gly His Phe Asp Thr Ala Thr Pro Pro Glu Arg
435      440      445
Met Leu Glu Ile Met Asn Arg Tyr Ala Arg Leu Gly Leu Pro Ile Ala
450      455      460
Ile Thr Glu Tyr Asp Phe Ala Thr Val Asp Glu Glu Leu Gln Ala Gln
465      470      475      480
Phe Thr Arg Asp Leu Met Ile Leu Ala Phe Ser His Pro Ala Val Ser
485      490      495

```


420 425 430
 Leu Ala Trp Thr His Ala Pro Gly Gln Ile Val Thr Ser Gly Trp Asn
 435 440 445
 Val Thr Ile Ala Gln Ser Gly Ser Ala Val Ser Ala Ser Asn Pro Ala
 450 455 460
 Gly Tyr Trp Asn Gly Val Ile Gly Ala Asn Gly Gly Lys Ile Ser Phe
 465 470 475 480
 Gly Phe Gln Gly Ser Leu Ala Gly Gly Ser Ala Val Ala Pro Thr Tyr
 485 490 495
 Phe Ala Leu Asn Gly Ala Ala Cys Asn Gly Ala Val Leu Pro Pro Thr
 500 505 510
 Ala Thr Phe Thr Pro Ser Pro Thr Ala Thr Met Cys Pro Gln Ala Thr
 515 520 525
 Pro Glu Leu Leu Val Val Gln Pro Val Thr Ser Pro Thr Thr Gln Leu
 530 535 540
 Ser Gln Thr Leu Val Val Arg Leu Gly Asn Gly Glu Trp Val Arg Ala
 545 550 555 560
 Ala Gly Pro Ala Gly Val Val Thr Val Thr Ala Pro Asp Pro Asp Gly
 565 570 575
 Tyr Phe Arg Leu Thr Ile Pro Leu Ala Asn Thr Ser Asn Ala Ile
 580 585 590
 Leu Val Glu Gly Arg Val Arg Val Ile Thr His Ser Asn Gly Cys Thr
 595 600 605
 Tyr Gly Gly Tyr Thr Leu Ser Arg Thr Val Thr Ile Val Gln Ala Ser
 610 615 620
 Ser Pro Val Thr Leu Thr Pro Thr Ala Thr Pro Ser Pro Thr Ala Thr
 625 630 635 640
 Ala Thr Pro Thr Val Thr Ala Thr Ser Pro Ser Gly Ala Cys Thr Val
 645 650 655
 Ala Tyr Ala Ile Thr Asn Asp Trp Gly Ser Gly Phe Thr Ala Asn Val
 660 665 670
 Thr Leu Thr Asn Thr Gly Gly Ser Ala Leu Asn Gly Trp Thr Leu Ala
 675 680 685
 Tyr Ala Phe Pro Gly Asn Gln Thr Ile Ser Asn Ala Trp Asn Gly Thr
 690 695 700
 Ala Val Gln Ser Gly Ser Ser Val Ser Val Thr Asn Ala Gly Trp Asn
 705 710 715 720
 Gly Ser Leu Pro Pro Asn Val Ser Ala Ser Phe Gly Phe Gln Ala Ser
 725 730 735
 Tyr Ser Gly Asn Asn Ser Val Pro Ala Ser Phe Thr Leu Asn Gly Ala
 740 745 750
 Leu Cys His
 755

<210> 331

<211> 1242

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 331

gtgttcaagg	gcttgcgcta	tttgcgttg	ctgtgcctga	gtgcgggact	ggtctttgcc	60
tgtgcgccac	ggctctgtgac	cgccccaccc	gatgggctaa	gcggggcaaat	taggctcctg	120
cgccaaggaa	ccctcactgt	ccttgctccag	aatgcccacg	ggcaacccat	tgccaacgcc	180
aagggtggtag	ctgctcagca	aaccatgccc	ttccccttg	gtgttgccct	agatacagca	240
atgtttgagc	cttccccgcc	acccgcagcc	aactggtacc	gcaacaccgc	tcgccaaaat	300
tttaatgccg	ctgtccatga	aaacgccctc	aagtggatg	cccttgaacc	ggagcagggc	360
aagctggact	ttacgatggc	ggatcgcatc	ctcgttgga	gtgaagccca	aggctggccg	420
atgcgggggc	acaccctctt	ttgggaagtt	gagcaattta	accccccatg	gctgaaaacg	480
ctgccaccag	agcaactgcg	ggctgccgtc	aagaaccatg	ccatgacggt	gtgtcgccat	540
taccgcgggc	gaatcaatga	atttgatgtc	aataatgaaa	tgctccacgg	taactttttc	600
cgcagtcgtt	tgggaaacgg	catagttaaa	gagatgttcg	agtgggtgccg	cgagggtaac	660
cccaggccg	tcctttatgt	gaacgactac	ggcattattg	agggcgatcg	cctcgacgac	720
tacgtgcagc	agattcgcg	tttactgggg	caagggttc	ccattggtgg	cattggcatt	780
caagccatt	tgggaatatcc	cttggatgca	gccaaagatga	aacgcgccct	tgataccctt	840
gcccaattca	acctgcccct	aaaaatcact	gaagttagtg	tcagccttgc	cgacgagcag	900
cagcaggcgg	agacactgcg	ccaaatctac	cgcattggtt	ttgcccatcc	agccgtcaaa	960

gagatcctcc	tgtggggatt	ttgggaaggc	aaccactggc	gaccccaagc	aggactgtac	1020
cgtcgcgact	tttccgcca	acctgctgcc	gaagcctatc	gacaactcct	ctttcaggag	1080
tggtggacca	ccagcaacgg	caaaactaat	gccgatgggc	gctggcagac	ccgcggctat	1140
gcggggcgct	atcgccctac	agtaacggcc	aacggccaga	ccattaaccg	cgacattgac	1200
ctaccagact	tggagagaac	cgtgaccgta	caattcccat	ga		1242

<210> 332
 <211> 413
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(28)

<400> 332

Met	Phe	Lys	Gly	Leu	Arg	Tyr	Leu	Leu	Leu	Leu	Cys	Leu	Ser	Ala	Gly
1				5				10					15		
Leu	Val	Phe	Ala	Cys	Ala	Pro	Arg	Ser	Val	Thr	Ala	Pro	Pro	Asp	Gly
			20					25				30			
Leu	Ser	Gly	Gln	Ile	Arg	Leu	Leu	Arg	Gln	Gly	Thr	Leu	Thr	Val	Leu
		35					40					45			
Val	Gln	Asn	Ala	Gln	Gly	Gln	Pro	Ile	Ala	Asn	Ala	Lys	Val	Val	Ala
	50					55					60				
Ala	Gln	Gln	Thr	His	Ala	Phe	Pro	Phe	Gly	Val	Ala	Leu	Asp	Thr	Ala
65					70				75						80
Met	Phe	Glu	Pro	Ser	Pro	Pro	Pro	Ala	Ala	Asn	Trp	Tyr	Arg	Asn	Thr
			85					90					95		
Ala	Arg	Gln	Asn	Phe	Asn	Ala	Ala	Val	His	Glu	Asn	Ala	Leu	Lys	Trp
			100					105					110		
Tyr	Ala	Leu	Glu	Pro	Glu	Gln	Gly	Lys	Leu	Asp	Phe	Thr	Met	Ala	Asp
		115					120					125			
Arg	Ile	Leu	Ala	Trp	Ser	Glu	Ala	Gln	Gly	Trp	Pro	Met	Arg	Gly	His
	130					135					140				
Thr	Leu	Phe	Trp	Glu	Val	Glu	Gln	Phe	Asn	Pro	Pro	Trp	Leu	Lys	Thr
145					150					155					160
Leu	Pro	Pro	Glu	Gln	Leu	Arg	Ala	Ala	Val	Lys	Asn	His	Ala	Met	Thr
			165						170					175	
Val	Cys	Arg	His	Tyr	Arg	Gly	Arg	Ile	Asn	Glu	Phe	Asp	Val	Asn	Asn
			180					185					190		
Glu	Met	Leu	His	Gly	Asn	Phe	Phe	Arg	Ser	Arg	Leu	Gly	Asn	Gly	Ile
	195						200					205			
Val	Lys	Glu	Met	Phe	Glu	Trp	Cys	Arg	Glu	Gly	Asn	Pro	Glu	Ala	Val
	210					215					220				
Leu	Tyr	Val	Asn	Asp	Tyr	Gly	Ile	Ile	Glu	Gly	Asp	Arg	Leu	Asp	Asp
225					230					235					240
Tyr	Val	Gln	Gln	Ile	Arg	Asp	Leu	Leu	Gly	Gln	Gly	Val	Pro	Ile	Gly
			245						250					255	
Gly	Ile	Gly	Ile	Gln	Ala	His	Leu	Glu	Tyr	Pro	Leu	Asp	Ala	Ala	Lys
			260					265					270		
Met	Lys	Arg	Ala	Leu	Asp	Thr	Leu	Ala	Gln	Phe	Asn	Leu	Pro	Leu	Lys
	275						280					285			
Ile	Thr	Glu	Val	Ser	Val	Ser	Leu	Ala	Asp	Glu	Gln	Gln	Gln	Ala	Glu
	290					295					300				
Thr	Leu	Arg	Gln	Ile	Tyr	Arg	Ile	Gly	Phe	Ala	His	Pro	Ala	Val	Lys
305					310					315					320
Glu	Ile	Leu	Leu	Trp	Gly	Phe	Trp	Glu	Gly	Asn	His	Trp	Arg	Pro	Gln
			325						330					335	
Ala	Gly	Leu	Tyr	Arg	Arg	Asp	Phe	Ser	Ala	Lys	Pro	Ala	Ala	Glu	Ala
			340					345					350		
Tyr	Arg	Gln	Leu	Leu	Phe	Gln	Glu	Trp	Trp	Thr	Thr	Ser	Asn	Gly	Lys
		355					360					365			
Thr	Asn	Ala	Asp	Gly	Arg	Trp	Gln	Thr	Arg	Gly	Tyr	Ala	Gly	Arg	Tyr
	370					375					380				
Arg	Leu	Thr	Val	Thr	Ala	Asn	Gly	Gln	Thr	Ile	Asn	Arg	Asp	Ile	Asp
385					390					395					400
Leu	Pro	Asp	Leu	Glu	Arg	Thr	Val	Thr	Val	Gln	Phe	Pro			

405

410

<210> 333
 <211> 1152
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 333
 atgaaaagac aattttattgg acgattgaga cttgtcacta tcctttcaat catagtgtatt 60
 atgggatgtg cttcaaaca aagtgtatcag aatgttgata acctaaagga cgccctcgac 120
 ggtttgttcc ttattggaac tgccatgaat acccccaga tcaccggaca ggataaccgg 180
 acgcttgaat tgatcaaaaa acacatgaac tccattgtgg cagaaaacgt tatgaaaagc 240
 ggactaatac agcccagcga aggggagttc gacttctcac ttgccgacca gtttgtgcaa 300
 ttcggtgttg acaacaacat gcacatcgta gggcataccc ttatctggca ttcgcaggct 360
 ccagggtggt tttttgtgga tgaaaacggt aatgatgtta gtcccgaagt tcttaagcaa 420
 aggatgaaag accacatcta cacagtagtt ggccgttaca aaggcaaagt gcacggttgg 480
 gatgtggtga atgaatgtat cgttgacgat ggggtcatggc gcaacagcaa gttttaccag 540
 atcctgggtg aagactttgt aaagtatgcc ttccagtttg cttcagaagc cgaccggaat 600
 gctgaattgt attacaacga ttattccatg gcaacttccc gccgcccga gggagtcgta 660
 aacatggtaa aaaatctaca ggcaaacggt attaaaattg acggaatagg aatgcagggc 720
 cacctgatga tcgaccatcc atcccttgaa gatttcgaaa ccagtttgct tgcctttgcc 780
 gatctgggtg tacatgttat gatcactgag cttgatgtat ctgtacttcc ttttcctacc 840
 cgcaacctcg gtgctgatgt atctctaaac atagcttaca acactgaact gaacccttat 900
 cccgatggat tgcctgatga tgtggcccaa aaacttcatt atcgttggtc cgatatatat 960
 cgtttatatta taaaacatca cgacaagatc acccgtgtta ctacctgggg tacagccgat 1020
 ggtatgtcat ggaagaacaa ctggccatt cgtggacgca cagactttcc tttattattc 1080
 gaccgcatg ttcaaccac accggtagta gctgatatta tcaaagaagc attggctgca 1140
 aagagaaaat ag 1152

<210> 334
 <211> 383
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(30)

<400> 334
 Met Lys Arg Gln Phe Ile Gly Arg Leu Arg Leu Val Thr Ile Leu Ser
 1 5 10 15
 Ile Ile Val Ile Met Gly Cys Ala Ser Asn Lys Ser Asp Gln Asn Val
 20 25 30
 Asp Asn Leu Lys Asp Ala Phe Asp Gly Leu Phe Leu Ile Gly Thr Ala
 35 40 45
 Met Asn Thr Pro Gln Ile Thr Gly Gln Asp Thr Arg Thr Leu Glu Leu
 50 55 60
 Ile Lys Lys His Met Asn Ser Ile Val Ala Glu Asn Val Met Lys Ser
 65 70 75 80
 Gly Leu Ile Gln Pro Ser Glu Gly Glu Phe Asp Phe Ser Leu Ala Asp
 85 90 95
 Gln Phe Val Gln Phe Gly Val Asp Asn Asn Met His Ile Val Gly His
 100 105 110
 Thr Leu Ile Trp His Ser Gln Ala Pro Gly Trp Phe Phe Val Asp Glu
 115 120 125
 Asn Gly Asn Asp Val Ser Pro Glu Val Leu Lys Gln Arg Met Lys Asp
 130 135 140
 His Ile Tyr Thr Val Val Gly Arg Tyr Lys Gly Lys Val His Gly Trp
 145 150 155 160
 Asp Val Val Asn Glu Cys Ile Val Asp Asp Gly Ser Trp Arg Asn Ser
 165 170 175
 Lys Phe Tyr Gln Ile Leu Gly Glu Asp Phe Val Lys Tyr Ala Phe Gln
 180 185 190
 Phe Ala Ser Glu Ala Asp Pro Asn Ala Glu Leu Tyr Tyr Asn Asp Tyr

```

      195              200              205
Ser Met Ala Leu Pro Gly Arg Arg Gln Gly Val Val Asn Met Val Lys
      210              215              220
Asn Leu Gln Ala Gln Gly Ile Lys Ile Asp Gly Ile Gly Met Gln Gly
      225              230              235
His Leu Met Ile Asp His Pro Ser Leu Glu Asp Phe Glu Thr Ser Leu
      245              250              255
Leu Ala Phe Ala Asp Leu Gly Val His Val Met Ile Thr Glu Leu Asp
      260              265              270
Val Ser Val Leu Pro Phe Pro Thr Arg Asn Leu Gly Ala Asp Val Ser
      275              280              285
Leu Asn Ile Ala Tyr Asn Thr Glu Leu Asn Pro Tyr Pro Asp Gly Leu
      290              295              300
Pro Asp Asp Val Ala Gln Lys Leu His Asp Arg Trp Leu Asp Ile Tyr
      305              310              315
Arg Leu Phe Ile Lys His His Asp Lys Ile Thr Arg Val Thr Thr Trp
      325              330              335
Gly Thr Ala Asp Gly Met Ser Trp Lys Asn Asn Trp Pro Ile Arg Gly
      340              345              350
Arg Thr Asp Phe Pro Leu Leu Phe Asp Arg Asp Phe Gln Pro Lys Pro
      355              360              365
Val Val Ala Asp Ile Ile Lys Glu Ala Leu Ala Ala Lys Arg Lys
      370              375              380

```

<210> 335
 <211> 849
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

```

<400> 335
atgattccaa ggatcgctcct ggccgtccgc atatccccta cttttctcag cccacaaaaa      60
ggggtaataa aaatgataaa gcgggctttt atgataaccc tggcggcctt cctcctcctt      120
ttcgccctaa attccctgcc tatccatgcc ggggccaag gcggggagga aaagtttacc      180
cccaagggtc tcgtggagca cgttttcgag aataacgact tccacgggtg ggtcccccg      240
ggcgggggtc ggaccatttc cattaccaat gaggcggccc atagcgggtc ctctgcctg      300
aagatcaccc gccggactca agcttgcat atgccgcggg tggagatcac caagtactta      360
gaaaagggag ctaagtataa gatcgaattg tacgtcaagc tccccgcggg cacctcgccg      420
cgcaagttcc agctggcggg tctcaccgt tatctcgaag gcaaccagac cagggacaaa      480
gaggactcca tctcggacga ggtggagggt accgccgata cctggaccaa ggtcgagggc      540
gagtacgtct tctgaccggc ggccatcggc gcctacgtct acccctacct caagggcgac      600
cccgaggggg cctatgcccc ctatctcatc gatgatttca agatcaccac gatcgcccc      660
gcccccaaga agaccgccgc taccgccgcg gcaaaagagg cagaagagcc cttaatcgag      720
accgatatac catccttaaa agacgtctgc gcgtcctact tcgagatcgg cgcggccatc      780
gagccatatg agttattctc caagccccac gatcagctgc tccggaaaca tttcaacacc      840
gttggttga

```

<210> 336
 <211> 282
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(50)

```

<400> 336
Met Ile Pro Arg Ile Val Leu Ala Val Arg Ile Ser Pro Thr Phe Leu
  1              5              10              15
Ser Pro Gln Lys Gly Val Ile Lys Met Ile Lys Arg Ala Phe Met Ile
      20              25              30
Thr Leu Ala Ala Phe Leu Leu Leu Phe Ala Leu Asn Ser Leu Pro Ile
      35              40              45
His Ala Gly Ala Glu Gly Gly Glu Glu Lys Phe Thr Pro Lys Val Ile
      50              55              60

```

Val Glu His Gly Phe Glu Asn Asn Asp Phe His Gly Trp Val Pro Arg
 65 70 75 80
 Gly Gly Val Gly Thr Ile Ser Ile Thr Asn Glu Ala Ala His Ser Gly
 85 90 95
 Ser Ser Cys Leu Lys Ile Thr Gly Arg Thr Gln Ala Trp His Met Pro
 100 105 110
 Arg Val Glu Ile Thr Lys Tyr Leu Glu Lys Gly Ala Lys Tyr Lys Ile
 115 120 125
 Glu Leu Tyr Val Lys Leu Pro Ala Gly Thr Ser Pro Arg Lys Phe Gln
 130 135 140
 Leu Ala Val Leu Thr Arg Tyr Leu Glu Gly Asn Gln Thr Arg Asp Lys
 145 150 155 160
 Glu Asp Ser Ile Ser Asp Glu Val Glu Val Thr Ala Asp Thr Trp Thr
 165 170 175
 Lys Val Glu Gly Glu Tyr Val Phe Asp Pro Ala Ala Ile Gly Ala Tyr
 180 185 190
 Val Tyr Pro Tyr Leu Lys Gly Asp Pro Ala Gly Ala Tyr Ala Pro Tyr
 195 200 205
 Leu Ile Asp Asp Phe Lys Ile Thr Thr Ile Ala Pro Ala Pro Lys Lys
 210 215 220
 Thr Ala Ala Thr Ala Ala Lys Glu Ala Glu Glu Pro Leu Ile Glu
 225 230 235 240
 Thr Asp Ile Pro Ser Leu Lys Asp Val Cys Ala Ser Tyr Phe Glu Ile
 245 250 255
 Gly Ala Ala Ile Glu Pro Tyr Glu Leu Phe Ser Lys Pro His Asp Gln
 260 265 270
 Leu Leu Arg Lys His Phe Asn Thr Val Gly
 275 280

<210> 337
 <211> 870
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 337
 atgaagcccg acagcgtgct ggatgtaaac gccagcaaaa agctctccgc ccaggatgaa 60
 accgccgtgg cggtgaaatt cgacgccgcc cgcgccctgc tggattttgt caaggaaaac 120
 gggctcaagg tgcacgggtca cgtgctggta tggcattccc agacgccgga agccttcttc 180
 cacgagggtc atgatgccgc caggccctac gtggggcggg acgtgatgct ggggcgcatg 240
 aaaaactaca tcaaggccgt gtttgaatac actgagacca attaccccg cgctatcgct 300
 ttcctgggacg tagtgaacga agccatcgac gacggcacca acaagctgcg ccagtccaac 360
 tggttcaaaa ccgttggcga ggatttcgtg ctccgcgcct ttgaatacgc caggaaatac 420
 gcccccgaag gcacgctgct ttattacaac gattacaaca ccgccatgcc cggcaagctg 480
 aacggcatcg ccaatctgct caaagccctc atcgccgagg gcaacatcga cggctacggc 540
 ttccaaatgc accacagcgt gggcttcccc tccatggaaa tgatttcgc gtctgtggag 600
 cgcacgcgcg gcatgggcct taagctccgg gtcagcgaat tggacgtggg caccgacgga 660
 aacaccgaaa gcagcttcac caagcaggcg gaaaaatag ccgccatcat gcggctgctg 720
 ctggattata aggatcaaat ggaagccgtg caggtatggg gcctcaccga cgatatgagc 780
 tggcgccggg ccaactatcc cctgctcttc gacggcaaat tcaaccccaa gcccgccctc 840
 tacgccgtgg ctgaccata cgcaaaaataa 870

<210> 338
 <211> 289
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 338
 Met Lys Pro Asp Ser Val Leu Asp Val Asn Ala Ser Lys Lys Leu Ser
 1 5 10 15
 Ala Gln Asp Glu Thr Ala Val Ala Val Lys Phe Asp Ala Ala Arg Ala
 20 25 30
 Leu Leu Asp Phe Val Lys Glu Asn Gly Leu Lys Val His Gly His Val
 35 40 45

Leu Val Trp His Ser Gln Thr Pro Glu Ala Phe Phe His Glu Gly Tyr
 50 55 60
 Asp Ala Ala Arg Pro Tyr Val Gly Arg Asp Val Met Leu Gly Arg Met
 65 70 75 80
 Lys Asn Tyr Ile Lys Ala Val Phe Glu Tyr Thr Glu Thr Asn Tyr Pro
 85 90 95
 Gly Val Ile Val Ser Trp Asp Val Val Asn Glu Ala Ile Asp Asp Gly
 100 105 110
 Thr Asn Lys Leu Arg Gln Ser Asn Trp Phe Lys Thr Val Gly Glu Asp
 115 120 125
 Phe Val Leu Arg Ala Phe Glu Tyr Ala Arg Lys Tyr Ala Pro Glu Gly
 130 135 140
 Thr Leu Leu Tyr Tyr Asn Asp Tyr Asn Thr Ala Met Pro Gly Lys Leu
 145 150 155 160
 Asn Gly Ile Ala Asn Leu Leu Lys Ala Leu Ile Ala Glu Gly Asn Ile
 165 170 175
 Asp Gly Tyr Gly Phe Gln Met His His Ser Val Gly Phe Pro Ser Met
 180 185 190
 Glu Met Ile Ser Ala Ser Val Glu Arg Ile Ala Gly Met Gly Leu Lys
 195 200 205
 Leu Arg Val Ser Glu Leu Asp Val Gly Thr Asp Gly Asn Thr Glu Ser
 210 215 220
 Ser Phe Thr Lys Gln Ala Glu Lys Tyr Ala Ala Ile Met Arg Leu Leu
 225 230 235 240
 Leu Asp Tyr Lys Asp Gln Met Glu Ala Val Gln Val Trp Gly Leu Thr
 245 250 255
 Asp Asp Met Ser Trp Arg Arg Ala Asn Tyr Pro Leu Leu Phe Asp Gly
 260 265 270
 Lys Phe Asn Pro Lys Pro Ala Phe Tyr Ala Val Ala Asp Pro Tyr Ala
 275 280 285
 Lys

<210> 339
 <211> 1125
 <212> DNA
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 339
 atgcctatgg agcgacccac tttcttgctg tttcttgctt tttttcttct ttttaccatg 60
 attttcgccg ccggagggtg ccgacccctt gcccttcac ggatggagat cgagacggat 120
 atccccctcc tcaaggaagt cgccgcttct tatttcgaga tcggcgcgcc cgctcgagccg 180
 tatcagttat cctctccacc ccacgatgcc cttctgcgga aacattttta ctgcctcgtg 240
 gcggagaacg tcatgaagcc cgcctccatc cagccatcgg aggggtatct caactggacc 300
 gaagcggaca agatcgtgaa ctacgccaaa gccacggga tgaagctccg cttccatacc 360
 ctctgtctggc ataatacagg cccggattgg ttcttcgctg gtaacgacaa aaccgcctt 420
 ttgcagcgtc tggagaatca tatccggact atcattaaaa gatatggcga taaggctcgc 480
 tattgggacg tgggtgaacga agtaatagac gacaacggcg gtatgcgaaa cagcaagtgg 540
 taccagatca ccgggaagga ctacatcaag accgccttc ggggtggcaga cgacgagctc 600
 aggaagaatg ggtggaggaa agaaggctgt cagctctata tcaacgacta caacacccat 660
 aacccaacga agagagaggg gatctggcgc ttgatccaag agctccgggc ggaagggatt 720
 cccgtcgacg gagtaggcca ccagacgcat atcaatatcg aatggccgcc cgtaagccag 780
 atcgtggaat cgtatccgct cttcggcgaa aaaggcctcg ataaccaggt gaccgagctg 840
 gatgtgagca tctatacgaa tgacaaggat tcacatggta gttatcaggc catcccgcag 900
 gaagtcttca tcaagcaggg taatcgctac aaggaactct ttgaagggct aaaaagtgtg 960
 aaaaactacc tcagcaacgt caccttctgg ggcattggcg acgatcatac ctggctgaac 1020
 cgttggccca tcgaacggcc cgatgctcct cttcctttcg atatctatct caaggccaag 1080
 ccggcgtatt gggggatcgt ggatgctttg aagctttcgc ggtga 1125

<210> 340
 <211> 374
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.
 Page 256

<221> SIGNAL

<222> (1)...(23)

<400> 340

```

Met Pro Met Glu Arg Pro Thr Phe Leu Arg Phe Leu Ala Phe Phe Leu
 1      5      10      15
Leu Phe Thr Met Ile Phe Ala Ala Gly Gly Cys Arg Pro Leu Ala Pro
 20      25      30
Ser Arg Met Glu Ile Glu Thr Asp Ile Pro Ser Leu Lys Glu Val Ala
 35      40      45
Ala Ser Tyr Phe Glu Ile Gly Ala Ala Val Glu Pro Tyr Gln Leu Ser
 50      55      60
Ser Pro Pro His Asp Ala Leu Leu Arg Lys His Phe Asn Cys Leu Val
 65      70      75      80
Ala Glu Asn Val Met Lys Pro Ala Ser Ile Gln Pro Ser Glu Gly Tyr
 85      90      95
Phe Asn Trp Thr Glu Ala Asp Lys Ile Val Asn Tyr Ala Lys Ala His
100      105      110
Gly Met Lys Leu Arg Phe His Thr Leu Val Trp His Asn Gln Val Pro
115      120      125
Asp Trp Phe Phe Ala Gly Asn Asp Lys Thr Arg Leu Leu Gln Arg Leu
130      135      140
Glu Asn His Ile Arg Thr Ile Ile Lys Arg Tyr Gly Asp Lys Val Asp
145      150      155      160
Tyr Trp Asp Val Val Asn Glu Val Ile Asp Asp Asn Gly Gly Met Arg
165      170      175      180
Asn Ser Lys Trp Tyr Gln Ile Thr Gly Lys Asp Tyr Ile Lys Thr Ala
185      190      195
Phe Arg Val Ala Asp Asp Glu Leu Arg Lys Asn Gly Trp Arg Lys Glu
200      205
Gly Arg Gln Leu Tyr Ile Asn Asp Tyr Asn Thr His Asn Pro Thr Lys
210      215      220
Arg Glu Gly Ile Trp Arg Leu Ile Gln Glu Leu Arg Ala Glu Gly Ile
225      230      235      240
Pro Val Asp Gly Val Gly His Gln Thr His Ile Asn Ile Glu Trp Pro
245      250      255
Pro Val Ser Gln Ile Val Glu Ser Ile Arg Phe Phe Gly Glu Lys Gly
260      265      270
Leu Asp Asn Gln Val Thr Glu Leu Asp Val Ser Ile Tyr Thr Asn Asp
275      280      285
Lys Asp Ser His Gly Ser Tyr Gln Ala Ile Pro Gln Glu Val Phe Ile
290      295      300
Lys Gln Gly Asn Arg Tyr Lys Glu Leu Phe Glu Gly Leu Lys Ser Val
305      310      315      320
Lys Asn Tyr Leu Ser Asn Val Thr Phe Trp Gly Met Ala Asp Asp His
325      330      335
Thr Trp Leu Asn Arg Trp Pro Ile Glu Arg Pro Asp Ala Pro Leu Pro
340      345      350
Phe Asp Ile Tyr Leu Lys Ala Lys Pro Ala Tyr Trp Gly Ile Val Asp
355      360      365
Ala Leu Lys Leu Ser Arg
370

```

<210> 341

<211> 1347

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 341

```

atgacaatta acaacaaaac tacagcgagt cctagttattc ccagcaccca caattccctc      60
ccgtcgcttc gcacactgtt taccaccagc ctgctcacgc tggccctgac cgctgcggt      120
ggttctttcca gcagcgacaa ggacccttca agctccagct ccagtgaatc atcaagtcc      180
agcgaatcct cgagctcagc ttccagcgaa tcctcgagca gtgagtccag cagtagctct      240
tccgcgggcc atttctccat cgagccggac ttccagctct acagcctggc caacttccc      300
gtgggcgtgg cggctctccgc cgccaacgag aacgacagca tcttcaacag tccgatgcc      360

```

```

gccgaacgct aggccgttat tattgagcac ttctctcagc tcaccgccgg caacatcatg 420
aaaatgagct acctgcagcc gagtcaaggc aacttcacct tcgatgacgc cgacgagttg 480
gttaacttcg cccaagccaa tggcatgacc gtacacggcc actccaccat ctggcacgcg 540
gactaccaag taccgaactt catgagaaac tttagaaggc accaggagga atgggcagaa 600
attctgaccg atcacgtcac taccatcatc gagcacttcc ccgacgatgt ggtcatcagc 660
tgggacgtgg tgaacgaggc tgtcgatcaa ggcacggcga acggctggcg ccattcggtg 720
ttctacaatg cattcgacgc cccggaagaa ggcgacattc ccgaatacat caaagtcgct 780
ttccgcgccc cgcgcgaggc tgacgccaac gtagacctct actacaacga ctacgacaat 840
accgccaatg cccagcgccct ggccaaaaca ctgcaaattg ccgaggtact ggacgccgaa 900
ggcaccattg acggcgctcg tttccagatg cacgcctaca tggattaccc gagcctgacc 960
cattttgaaa acgccttccg gcaagtcgtc gacctggggc tcaaagtga agttaccgag 1020
ctggacgtat ccgtagtcaa cccctacggc ggcgaaagcac ctccacaacc ggaatacgac 1080
aaagaactgg ccggcgcgca aaaactgcgc ttctgccaaa tcgccgaagt ttacatgaac 1140
actgtacccg aggagttacg cggtggcttc accgtctggg gcctgaccga tgatgaaagt 1200
tggctgatgc aacagttcag aaacgccacc ggcgccgact acgacgacgt ctggccgtta 1260
ctgttcaatg ccgacaaatc cgccaaaccg gcactgcaag gcgtggccga cgcttttacc 1320
ggacaaacct gcacctccga gttctaa 1347

```

<210> 342
 <211> 448
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(45)

<400> 342
 Met Thr Ile Asn Asn Lys Thr Thr Ala Ser Pro Ser Ile Pro Ser Thr
 1 5 10 15
 His Asn Ser Leu Pro Ser Leu Arg Thr Leu Phe Thr Thr Ser Leu Leu
 20 25 30
 Thr Leu Ala Leu Thr Ala Cys Gly Gly Ser Ser Ser Ser Asp Lys Asp
 35 40 45
 Pro Ser Ser Ser Ser Ser Ser Glu Ser Ser Ser Ser Ser Glu Ser Ser
 50 55 60
 Ser Ser Ala Ser Ser Glu Ser Ser Ser Ser Glu Ser Ser Ser Ser Ser
 65 70 75 80
 Ser Ala Gly His Phe Ser Ile Glu Pro Asp Phe Gln Leu Tyr Ser Leu
 85 90 95
 Ala Asn Phe Pro Val Gly Val Ala Val Ser Ala Ala Asn Glu Asn Asp
 100 105 110
 Ser Ile Phe Asn Ser Pro Asp Ala Ala Glu Arg Gln Ala Val Ile Ile
 115 120 125
 Glu His Phe Ser Gln Leu Thr Ala Gly Asn Ile Met Lys Met Ser Tyr
 130 135 140
 Leu Gln Pro Ser Gln Gly Asn Phe Thr Phe Asp Asp Ala Asp Glu Leu
 145 150 155 160
 Val Asn Phe Ala Gln Ala Asn Gly Met Thr Val His Gly His Ser Thr
 165 170 175
 Ile Trp His Ala Asp Tyr Gln Val Pro Phe Met Arg Asn Phe Glu
 180 185 190
 Gly Asp Gln Glu Glu Trp Ala Glu Ile Leu Thr Asp His Val Thr Thr
 195 200 205
 Ile Ile Glu His Phe Pro Asp Val Val Ile Ser Trp Asp Val Val
 210 215 220
 Asn Glu Ala Val Asp Gln Gly Thr Ala Asn Gly Trp Arg His Ser Val
 225 230 235 240
 Phe Tyr Asn Ala Phe Asp Ala Pro Glu Glu Gly Asp Ile Pro Glu Tyr
 245 250 255
 Ile Lys Val Ala Phe Arg Ala Ala Arg Glu Ala Asp Ala Asn Val Asp
 260 265 270
 Leu Tyr Tyr Asn Asp Tyr Asp Asn Thr Ala Asn Ala Gln Arg Leu Ala
 275 280 285
 Lys Thr Leu Gln Ile Ala Glu Val Leu Asp Ala Glu Gly Thr Ile Asp
 290 295 300
 Gly Val Gly Phe Gln Met His Ala Tyr Met Asp Tyr Pro Ser Leu Thr

305	His	Phe	Glu	Asn	Ala	310	Phe	Arg	Gln	Val	Val	315	Asp	Leu	Gly	Leu	Lys	320	Val
					325							330					335		
	Lys	Val	Thr	Glu	Leu	Asp	Val	Ser	Val	Val	Val	Asn	Pro	Tyr	Gly	Gly	Glu		
			340						345						350				
	Ala	Pro	Pro	Gln	Pro	Glu	Tyr	Asp	Lys	Glu	Leu	Ala	Gly	Ala	Gln	Lys			
			355					360						365					
	Leu	Arg	Phe	Cys	Gln	Ile	Ala	Glu	Val	Tyr	Met	Asn	Thr	Val	Pro	Glu			
		370				375						380							
	Glu	Leu	Arg	Gly	Gly	Phe	Thr	Val	Trp	Gly	Leu	Thr	Asp	Asp	Glu	Ser			
	385					390					395					400			
	Trp	Leu	Met	Gln	Gln	Phe	Arg	Asn	Ala	Thr	Gly	Ala	Asp	Tyr	Asp	Asp			
				405						410					415				
	Val	Trp	Pro	Leu	Phe	Asn	Ala	Asp	Lys	Ser	Ala	Lys	Pro	Ala	Leu				
			420					425					430						
	Gln	Gly	Val	Ala	Asp	Ala	Phe	Thr	Gly	Gln	Thr	Cys	Thr	Ser	Glu	Phe			
		435					440					445							

<210> 343
 <211> 2217
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 343																			
atggttgagc	acgaagctga	gcttcatgat	taccgcgaac	gaatatgcga	agttgcttgg														60
caattcgctg	ggcctggagg	gcgagaaggc	aaagcggatt	ttgcgggaaa	tcaacaactg														120
gctgttttga	agatgaaatt	atcgagaggg	agggggcatg	gcgttttgaa	acgagcaggt														180
ctgattgggc	ttattgcagc	gatattgggc	tgcgcggcgc	tgcttatgca	caatgagatc														240
tcacccctgg	gaattcagct	gagatcatgg	ctcagaggga	gcgacaatgt	gaacgctagc														300
tgggaaaagg	attggaagac	ggctgctaac	gagcaaatac	agcagctccg	caagcgcaat														360
gtggagatcg	aggtcgctga	tctgaacgga	aaccgcgtgc	ctggggctac	cgttcgcgcg														420
gttcagcgca	cgcatacgtt	tggcttcggc	accgccatca	accgaacggc	gttgagcaat														480
ccggtgtacg	ccgattttgt	caaaaaccgt	ttcgaatggg	tgaccttcga	gaacgaggcc														540
aagtggctct	ggaatgaggg	cgtacaaggg	cgggtctatt	atcgggaggc	cgatcagctg														600
ctcgaaattg	ccaggcaaaa	cgggctgaag	gtgcgcggac	ataatctggt	ctgggaggcg														660
gagaaatatc	agccgcagtg	ggtgaagagt	ctgacgggcg	ctgcgctgaa	ggaagcgatc														720
gataaccggc	tgaacagcgc	cgtcctgcat	tttaagggca	attttctgca	ctgggacgtc														780
aacaacgaaa	tgtttcacgg	cagcttcttc	aaggatcgcc	tgggggaaga	aatctggacc														840
tatatgtata	agcgaacccg	ggaactcgat	ccgggcgtca	agctgttcgt	caacgattac														900
aattttatcg	agtaccgcc	ggaagcggat	tataaccagg	tcatccaagc	gctcatcgat														960
cgggggatgc	cgattgacgg	catcggcggc	caagggcatt	ttaacggagt	catcgatccc														1020
ttgttcgtta	agggaagact	ggataagctg	gctgagctga	atctgccgat	ctggattacc														1080
gaattcgatt	ccacgcataa	ggacgagaga	gtccgtgccg	ataatctgga	gaagatgtat														1140
cggtggcgt	tcgcccattc	ggcggctgaa	gggattgtca	tgtggggctt	ctgggcgggc														1200
tccattgga	agggcaactg	cggcgcgatc	gtgaatcaag	actggacgct	caatgccgcc														1260
ggacagcgat	accagcagct	tatggatgaa	tggacgacgg	tcgtcgaagg	cacgaccgat														1320
cagcgcgca	tgttttcggt	ccgggggttc	cacggaacct	acgatatgct	ggtcgattac														1380
cctggagcgg	cggctgtgaa	gcagtccttt	accttgagc	ccggctcttg	caatgcgaag														1440
ctgcacattc	cgttcgacgt	tcaggacaag	tccatcccg	aggctcctgc	caagctcagc														1500
gccgtgccc	cggattccca	ggttatgctg	agctggagca	agggtcaacg	ggcaaccggc														1560
tatacggtta	aaagcgcggt	cagcgccgac	ggtccctata	cgccgattgc	ccatcagctg														1620
ctcaccgaga	ccttcacgca	catcggtcta	gtgaaccgga	aagattatta	ttacgtgggtg														1680
agcgccagca	accattatag	tgagagccc	gattccgccc	cgatccgggc	cactccgcgt														1740
gccgcgggcg	agttacaac	gaatctcggt	cttcagttac	gctccgctga	tggagataac														1800
aactatcaaa	tgaagcctca	gttcacgata	aagaacgcag	gcaaagtgcc	catcccgtta														1860
agcgagctga	cgatccgcta	ctatttcacg	ccggagagca	cgcagccggt	ggataaccagg														1920
atcgactggg	cccaattcgg	agcagagcat	gtccagacga	cggctcgttc	gccatccgat														1980
gccgcggcgc	agccctatgt	cgagctcagc	ttcctggagt	cggcaggggc	catcccttcc														2040
gatacgacat	taggcaatat	tcagctgcgc	atctttaaca	gcgatggctc	ttcgttcgat														2100
aaaacgaacg	attattcctt	cgacccgacg	aaaaaggctt	atacggcggtg	ggagaagggtc														2160
acgctttatc	ggaacgggga	actggtttgg	gggatagagc	cttggggcgc	gaagtaa														2217

<210> 344
 <211> 738
 <212> PRT
 <213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 344

```

Met Val Glu His Glu Ala Glu Leu His Asp Tyr Arg Glu Arg Ile Cys
 1      5      10      15
Glu Val Ala Trp Gln Phe Ala Gly Pro Gly Gly Arg Glu Gly Lys Ala
 20      25      30
Asp Phe Ala Gly Asn Gln Gln Leu Val Leu Lys Met Lys Leu Ser
 35      40      45
Arg Gly Arg Gly His Gly Val Leu Lys Arg Ala Gly Leu Ile Gly Leu
 50      55      60
Ile Ala Ala Ile Leu Gly Cys Ala Ala Leu Leu Met His Asn Glu Ile
 65      70      75      80
Ser Ser Leu Gly Ile Gln Leu Arg Ser Trp Leu Arg Gly Ser Asp Asn
 85      90      95
Val Asn Ala Ser Trp Glu Lys Asp Trp Lys Thr Ala Ala Asn Glu Gln
100      105      110
Ile Glu Gln Leu Arg Lys Arg Asn Val Glu Ile Glu Val Val Asp Leu
115      120      125
Asn Gly Asn Pro Leu Pro Gly Ala Thr Val Arg Ala Val Gln Arg Thr
130      135      140
His Gln Phe Gly Phe Gly Thr Ala Ile Asn Arg Thr Ala Leu Ser Asn
145      150      155      160
Pro Val Tyr Ala Asp Phe Val Lys Asn Arg Phe Glu Trp Val Thr Phe
165      170      175
Glu Asn Glu Ala Lys Trp Leu Trp Asn Glu Ala Val Gln Gly Arg Val
180      185      190
Tyr Tyr Arg Glu Ala Asp Gln Leu Leu Glu Phe Ala Arg Gln Asn Gly
195      200      205
Leu Lys Val Arg Gly His Asn Leu Phe Trp Glu Ala Glu Lys Tyr Gln
210      215      220
Pro Gln Trp Val Lys Ser Leu Thr Gly Ala Ala Leu Lys Glu Ala Ile
225      230      235      240
Asp Asn Arg Leu Asn Ser Ala Val Leu His Phe Lys Gly Asn Phe Leu
245      250      255
His Trp Asp Val Asn Asn Glu Met Phe His Gly Ser Phe Phe Lys Asp
260      265      270
Arg Leu Gly Glu Glu Ile Trp Thr Tyr Met Tyr Lys Arg Thr Arg Glu
275      280      285
Leu Asp Pro Gly Val Lys Leu Phe Val Asn Asp Tyr Asn Phe Ile Glu
290      295      300
Tyr Pro Pro Glu Arg Asp Tyr Asn Gln Val Ile Gln Ala Leu Ile Asp
305      310      315      320
Arg Gly Met Pro Ile Asp Gly Ile Gly Ala Gln Gly His Phe Asn Gly
325      330      335
Val Ile Asp Pro Leu Phe Val Lys Gly Arg Leu Asp Lys Leu Ala Glu
340      345      350
Leu Asn Leu Pro Ile Trp Ile Thr Glu Phe Asp Ser Thr His Lys Asp
355      360      365
Glu Arg Val Arg Ala Asp Asn Leu Glu Lys Met Tyr Arg Leu Ala Phe
370      375      380
Ala His Pro Ala Val Glu Gly Ile Val Met Trp Gly Phe Trp Ala Gly
385      390      395      400
Ser His Trp Lys Gly Thr Asp Gly Ala Ile Val Asn Gln Asp Trp Thr
405      410      415
Leu Asn Ala Ala Gly Gln Arg Tyr Gln Gln Leu Met Asp Glu Trp Thr
420      425      430
Thr Val Val Glu Gly Thr Thr Asp Gln Arg Gly Met Phe Ser Phe Arg
435      440      445
Gly Phe His Gly Thr Tyr Asp Met Leu Val Asp Tyr Pro Gly Ala Ala
450      455      460
Ala Val Lys Gln Ser Phe Thr Leu Glu Pro Gly Ser Gly Asn Ala Lys
465      470      475      480
Leu His Ile Pro Phe Asp Val Gln Asp Lys Ser Ile Pro Glu Ala Pro
485      490      495
Ala Lys Leu Ser Ala Ala Ala Ala Asp Ser Gln Val Met Leu Ser Trp
500      505      510

```

Ser Lys Val Asn Gly Ala Thr Gly Tyr Thr Val Lys Ser Ala Val Ser
 515 520 525
 Ala Asp Gly Pro Tyr Thr Pro Ile Ala His Gln Leu Leu Thr Glu Thr
 530 535 540
 Phe Thr His Ile Gly Leu Val Asn Arg Lys Asp Tyr Tyr Tyr Val Val
 545 550 555 560
 Ser Ala Ser Asn His Leu Gly Glu Ser Pro Asp Ser Ala Pro Ile Arg
 565 570 575
 Ala Thr Pro Arg Ala Ala Gly Glu Leu Gln Thr Asn Leu Val Leu Gln
 580 585 590
 Tyr Arg Ser Ala Asp Gly Asp Asn Tyr Gln Met Lys Pro Gln Phe
 595 600 605
 Thr Ile Lys Asn Ala Gly Lys Val Pro Ile Pro Leu Ser Glu Leu Thr
 610 615 620
 Ile Arg Tyr Tyr Phe Thr Pro Glu Ser Thr Gln Pro Val Asp Thr Arg
 625 630 635 640
 Ile Asp Trp Ala Gln Phe Gly Ala Glu His Val Gln Thr Thr Val Val
 645 650 655
 Pro Pro Ser Asp Ala Ala Ala His Ala Tyr Val Glu Leu Ser Phe Leu
 660 665 670
 Glu Ser Ala Gly Ala Ile Pro Ser Asp Thr Thr Leu Gly Asn Ile Gln
 675 680 685
 Leu Arg Ile Phe Asn Ser Asp Gly Ser Ser Phe Asp Lys Thr Asn Asp
 690 700
 Tyr Ser Phe Asp Pro Thr Lys Lys Ala Tyr Thr Ala Trp Glu Lys Val
 705 710 715 720
 Thr Leu Tyr Arg Asn Gly Glu Leu Val Trp Gly Ile Glu Pro Trp Gly
 725 730 735
 Ala Lys

<210> 345

<211> 849

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 345

atgaagatga	cctacatgca	tccggctgaa	gatacttact	cgtttggtca	agcggatcag	60
ttgggtcaact	gggcgaaagc	gaatggtatt	ggcgtgcacg	gccacactct	ggtttggcac	120
tccgaataacc	aggtacccaa	ttggatgaaa	aattactctg	gtgatgcaac	tgcatcccaa	180
accatgctca	acacccatgt	gaaaactgtg	gctgagcatt	ttgctggcga	actggacagc	240
tgggacgttg	tgaatgaagt	gctggagccg	ggctccaatg	gttgctggcg	tgaaaactct	300
ctgtttctacc	agaagcttgg	caaagacttt	gtcgcgaacg	cattccgtgc	agctcgcgag	360
ggcgatccca	atgcagactt	gtattacaac	gattactcga	ctgaaaatgg	tgtaacttcc	420
gatgagaagt	tcagttgttt	gttggaaacta	gtcgtgagc	ttctggaagc	ggacgtgccg	480
attacaggtg	ttggtttcca	aatgcacgtg	caggcgacgt	ggcctagcaa	tgccaacatc	540
ggcaaggcat	tcaaagccat	cgcggatcgc	ggtctgaaag	ttaaaatttc	tgagctcgat	600
gttcctgtta	acaaccctta	cggaaccact	aattttccgc	aatacagcag	ttttaccgcg	660
gaagccgccg	agctgcagaa	gcagcgctac	aagggcatta	tgcaagcgta	ccttgataac	720
gtaccggcca	acctgcgtgg	tggtttcacc	gtgtggggcg	tttgggatgg	cgatagctgg	780
atcatgacgt	tcagccagta	caccaacgct	aacgccaacg	actggccact	gttggttcacc	840
gggccgtag						849

<210> 346

<211> 282

<212> PRT

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 346

Met Lys Met Thr Tyr Met His Pro Ala Glu Asp Thr Tyr Ser Phe Gly	
1 5 10 15	
Gln Ala Asp Gln Leu Val Asn Trp Ala Lys Ala Asn Gly Ile Gly Val	
20 25 30	

His Gly His Thr Leu Val Trp His Ser Glu Tyr Gln Val Pro Asn Trp
 35 40 45
 Met Lys Asn Tyr Ser Gly Asp Ala Thr Ala Phe Gln Thr Met Leu Asn
 50 55 60
 Thr His Val Lys Thr Val Ala Glu His Phe Ala Gly Glu Leu Asp Ser
 65 70 75 80
 Trp Asp Val Val Asn Glu Val Leu Glu Pro Gly Ser Asn Gly Cys Trp
 85 90 95
 Arg Glu Asn Ser Leu Phe Tyr Gln Lys Leu Gly Lys Asp Phe Val Ala
 100 105 110
 Asn Ala Phe Arg Ala Ala Arg Glu Gly Asp Pro Asn Ala Asp Leu Tyr
 115 120 125
 Tyr Asn Asp Tyr Ser Thr Glu Asn Gly Val Thr Ser Asp Glu Lys Phe
 130 135 140
 Ser Cys Leu Leu Glu Leu Val Asp Glu Leu Leu Glu Ala Asp Val Pro
 145 150 155 160
 Ile Thr Gly Val Gly Phe Gln Met His Val Gln Ala Thr Trp Pro Ser
 165 170 175
 Asn Ala Asn Ile Gly Lys Ala Phe Lys Ala Ile Ala Asp Arg Gly Leu
 180 185 190
 Lys Val Lys Ile Ser Glu Leu Asp Val Pro Val Asn Asn Pro Tyr Gly
 195 200 205
 Thr Thr Asn Phe Pro Gln Tyr Ser Ser Phe Thr Ala Glu Ala Ala Glu
 210 215 220
 Leu Gln Lys Gln Arg Tyr Lys Gly Ile Met Gln Ala Tyr Leu Asp Asn
 225 230 235 240
 Val Pro Ala Asn Leu Arg Gly Gly Phe Thr Val Trp Gly Val Trp Asp
 245 250 255
 Gly Asp Ser Trp Ile Met Thr Phe Ser Gln Tyr Thr Asn Ala Asn Ala
 260 265 270
 Asn Asp Trp Pro Leu Leu Phe Thr Gly Pro
 275 280

<210> 347
 <211> 1794
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 347
 atgccggttt tgttcgccct gtttcttgtt gcctcgtcct gcgcggcgca gtcgctggcc 60
 gggccggttt ccctgcttgg cgggagatgcg ggcgcggcgt tccgctatac cgggccatcg 120
 gcggggcgcg cgagcggctc ggccgaatgg gtggcggtgg agaactatgcc gttcacgcac 180
 gcctggcgcg tgcgcacgaa tccgctgccc gagagcgcg gcaacgaatg ggacctgcgc 240
 atccgcgccc gggagcgggc ggctgtttcg gcaggggaca agatcctggc cgagttcttg 300
 atgcgctgcg tggagcccga aaacggcgac tgcattctgc gcctgaacgt ggagcgcgac 360
 gggctgcgct ggaccaaacc catcagcaac ccctaccggc tgggcccggg gtggcgcgcg 420
 ttccgcgtgc tgttcgagat gcgggagagc tacgcgcgcg gcggctacat gatcgatttc 480
 tggatgggccc agcaggtgca gacggcgga gtggcgggg tttccctgct gaattacggt 540
 ccgcaggcca cggccgagca gcttggcctg gaccggttt atgaggcgcg ggcggcgagc 600
 gccgcgtggc ggcaggcggc cgagcagcgg atcgaggaga tccggaaagc gggcatgatc 660
 atcgtggcgg tgacgccgga cggcgagccg atcgaggcg ctgaaatccg ggcgaagctg 720
 aagcggcacg cgttcgggtg gggcacggct gtggcgccat caccgcttct ggggacggga 780
 acggacagcg agcgtaccg caacttcac cgcgagaact tcaacatggc ggtgctcgag 840
 aacgacctga aatggggccc gttcgaagag aaccgcaacc gcgcgatgaa cgcgctgcgc 900
 tggctgcatg agaacgggat cacgtggatc cgcgggcaca atctcgtctg gccgggctgg 960
 cggttgatgc cgaacgacgt gcgcaacctg gcgaacaatc ccgaggcgct gcggcagcgg 1020
 attctggacc gcatccggga cacggccacg gccacgcgcg ggctgggtgt gactggggac 1080
 gtcgtcaacg agccggtggc cgagcgcgac gtgctgaaca ttctgggcga cgaggtgatg 1140
 gcggatggt tccgcgccgc gaaggagtgc gatcccagg cgaggatgtt catcaatgag 1200
 tacgacattc tggcggcgaa cggggccaat ctgcggaagc agaacgcgta ttaccgcatg 1260
 atcgagatgc tgttgaagct cgaggcgccc gtggagggca tcggcttcca gggccacttc 1320
 gacacggcca cgccgcccga gcggatgctg gagatcatga accggtacgc ccggctcggg 1380
 ctgccgatcg ccatcaccga gtacgatttc gccacggcgg acgaggagct gcaggcgagc 1440
 ttcacgcgcg acctgatgat tctcgcttc agccatccgg cggtttcgga cttcctgatg 1500
 tggggcttct ggggaaggag ccactggaag ccgctggcg ccatgatccg gcgcgactgg 1560
 agcgagaagc cgatgtaccg cgtctggcgc gagctgattc tcgagcgctg gcagacggat 1620

gaaacaggcg	tgacgccgga	gcacgggtgcc	atctacgtgc	ggggcttcaa	gggcgactac	1680
gagatcacgg	tgaaggcggg	cgggcaggaa	gtccgggtgc	cgtaacacgt	gaaagaagac	1740
ggccaggtgc	tgtgggtgac	ggtgggcggg	gcttctgaag	agcgcgtgca	gtaa	1794

<210> 348
 <211> 597
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(20)

<400> 348
 Met Pro Val Leu Phe Ala Leu Phe Leu Val Ala Ser Ser Cys Ala Ala
 1 5 10 15
 Gln Ser Leu Ala Gly Pro Val Ser Leu Leu Gly Gly Asp Ala Gly Ala
 20 25 30
 Ala Phe Arg Tyr Thr Gly Pro Ser Ala Gly Ala Ala Ser Gly Ser Ala
 35 40 45
 Glu Trp Val Ala Val Glu Asn Met Pro Phe Thr His Ala Trp Arg Leu
 50 55 60
 Arg Thr Asn Pro Leu Pro Glu Ser Gly Gly Asn Glu Trp Asp Leu Arg
 65 70 75 80
 Ile Arg Ala Arg Gly Ala Ala Ala Val Ser Ala Gly Asp Lys Ile Leu
 85 90 95
 Ala Glu Phe Trp Met Arg Cys Val Glu Pro Glu Asn Gly Asp Cys Ile
 100 105 110
 Leu Arg Leu Asn Val Glu Arg Asp Gly Ser Pro Trp Thr Lys Ser Ile
 115 120 125
 Ser Asn Pro Tyr Pro Val Gly Arg Glu Trp Arg Arg Phe Arg Val Leu
 130 135 140
 Phe Glu Met Arg Glu Ser Tyr Ala Ala Gly Gly Tyr Met Ile Asp Phe
 145 150 155 160
 Trp Met Gly Gln Gln Val Gln Thr Ala Glu Val Gly Gly Ile Ser Leu
 165 170 175
 Leu Asn Tyr Gly Pro Gln Ala Thr Ala Glu Gln Leu Gly Leu Asp Arg
 180 185 190
 Phe Tyr Glu Gly Ala Ala Ala Asp Ala Ala Trp Arg Gln Ala Ala Glu
 195 200 205
 Gln Arg Ile Glu Glu Ile Arg Lys Ala Gly Met Ile Ile Val Ala Val
 210 215 220
 Thr Pro Asp Gly Glu Pro Ile Glu Gly Ala Glu Ile Arg Ala Lys Leu
 225 230 235 240
 Lys Arg His Ala Phe Gly Trp Gly Thr Ala Val Ala Ala Ser Arg Leu
 245 250 255
 Leu Gly Thr Gly Thr Asp Ser Glu Arg Tyr Arg Asn Phe Ile Arg Glu
 260 265 270
 Asn Phe Asn Met Ala Val Leu Glu Asn Asp Leu Lys Trp Gly Pro Phe
 275 280 285
 Glu Glu Asn Arg Asn Arg Ala Met Asn Ala Leu Arg Trp Leu His Glu
 290 295 300
 Asn Gly Ile Thr Trp Ile Arg Gly His Asn Leu Val Trp Pro Gly Trp
 305 310 315 320
 Arg Trp Met Pro Asn Asp Val Arg Asn Leu Ala Asn Asn Pro Glu Ala
 325 330 335
 Leu Arg Gln Arg Ile Leu Asp Arg Ile Arg Asp Thr Ala Thr Ala Thr
 340 345 350
 Arg Gly Leu Val Val His Trp Asp Val Val Asn Glu Pro Val Ala Glu
 355 360 365
 Arg Asp Val Leu Asn Ile Leu Gly Asp Glu Val Met Ala Asp Trp Phe
 370 375 380
 Arg Ala Ala Lys Glu Cys Asp Pro Glu Ala Arg Met Phe Ile Asn Glu
 385 390 395 400
 Tyr Asp Ile Leu Ala Ala Asn Gly Ala Asn Leu Arg Lys Gln Asn Ala
 405 410 415
 Tyr Tyr Arg Met Ile Glu Met Leu Leu Lys Leu Glu Ala Pro Val Glu

420 425 430
 Gly Ile Gly Phe Gln Gly His Phe Asp Thr Ala Thr Pro Pro Glu Arg
 435 440 445
 Met Leu Glu Ile Met Asn Arg Tyr Ala Arg Leu Gly Leu Pro Ile Ala
 450 455 460
 Ile Thr Glu Tyr Asp Phe Ala Thr Ala Asp Glu Glu Leu Gln Ala Gln
 465 470 475 480
 Phe Thr Arg Asp Leu Met Ile Leu Ala Phe Ser His Pro Ala Val Ser
 485 490 495
 Asp Phe Leu Met Trp Gly Phe Trp Glu Gly Ser His Trp Lys Pro Leu
 500 505 510
 Gly Ala Met Ile Arg Arg Asp Trp Ser Glu Lys Pro Met Tyr Arg Val
 515 520 525
 Trp Arg Glu Leu Ile Phe Glu Arg Trp Gln Thr Asp Glu Thr Gly Val
 530 535 540
 Thr Pro Glu His Gly Ala Ile Tyr Val Arg Gly Phe Lys Gly Asp Tyr
 545 550 555 560
 Glu Ile Thr Val Lys Ala Gly Gly Gln Glu Val Arg Val Pro Tyr Thr
 565 570 575
 Leu Lys Glu Asp Gly Gln Val Leu Trp Val Thr Val Gly Gly Ala Ser
 580 585 590
 Glu Glu Arg Val Gln
 595

<210> 349
 <211> 1794
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 349
 atgccggttt tgttcgccct gtttcttggt gcctcgtcct gcgcggcgca gtcgctggcc 60
 gggccggttt ccctgcttgg cggagatgcg ggcgcggcgt tccgctatac cgggccatcg 120
 gcgggcgcg cgagcggctc ggccgaatgg gtggcggtgg agaacatgcc gttcacgcac 180
 gcctggcggc tgcgcacgaa tccgctgccc gagagcggcg gcaacgaatg ggacctgcgc 240
 atccgcgccc gcggagcggc ggctgtttcg gcaggggaca agatcctggc cgagtcttg 300
 atgctgctcg tggagcccga aaacggcgac tgcattctgc gcctgaacgt ggagcgcgac 360
 gggctgcctg ggaccaaatt catcagcaac ccctaccggt tgggcccggg gtggcggcgg 420
 ttccgcgtgc tgttcgagat gcgggagagc tacgccggcg gcggctacat gatcgatttc 480
 tggatggggc agcaggtgca gacggcggaa gtgggcccga ttccctgct gaattacggt 540
 ccgcaggcca cggccgagca gcttggcctg gaccggttct atgaaggcgc ggcggcggac 600
 gccgcgtggc ggcaggcggc cgagcagcgg atcgaggaga tccggaaagc gggcatgatc 660
 atcgtggcgg tgacgcggga cggcgagcgg atcgaaggcg ccgagatccg ggcgaagctg 720
 aagcggcagc cgttcgggtg gggcacggcg gtggcggcat cacggcttct ggggacggga 780
 acggacagcg agcgtaccg caacttcac cgcgagaact tcaacatggc ggtgctcgag 840
 aacgacctga aatggggccc gttcgaggag aaccgcgccc gcgcaatgaa cgcgctgcgc 900
 tggctgcatg agaacgggat cacgtggatc cgcgggcaca atctcgtctg gccaggctgg 960
 cgggtgatgc cgagcgacgt gcgcaacctg gcgaacaatc ccgaagccct gcggcagcgg 1020
 attctggacc gcatccggga cacggccacc gccacgcgcg ggctggtcgt gactggggac 1080
 gtcgtcaacg agccggtggc cgagcgcgac gtgctgaaca ttctgggcga cgaggtgatg 1140
 gcggactggt tccgcgccgc cggggccaac ctgcggaagc agaacgcgta ctaccgcatg 1200
 tacgacattc tggcgccgaa cgaggcggcg gtagagggca tcggcttcca gggccatttc 1260
 atcgagatgc tgttgaagct gacccgctga gcggatgctg gagatcatga accggtacgc ccggtcggg 1320
 gacacggcta ccattaccga gtacgatttc gccacggtag acgaagagct gcaggcgacg 1380
 ctgccgatcg acctgatgat tctcgttttc agccatccgg cggtttcgga cttcttgatg 1440
 ttcacgcgcg ggggaagggag cacttgggag ccgctgggcg ccatgatccg gcgcgactgg 1500
 tggggcttct ggggaagggag cgtctggcgc gagctgatct tcgagcgctg gcagacggat 1560
 agcgagaagc gacgcccggg tgacgcccga gcacggggcc atctacgtgc ggggcttcaa gggcgactac 1620
 gaaacgggag gaaatcacgg tgaaggctgg cgggcaggaa gtccgggtgc cgtaacgcgt gaaagaagac 1680
 ggccaggtgc tgtgggtgac ggtgggcggg acttctgaag agcaggcgcc gtaa 1740
 1794

<210> 350
 <211> 597
 <212> PRT
 <213> Unknown

cataatggcg	caacagcggg	tgcaagggtg	gtttggcacg	aagatgctat	ttatatattta	1800
gccaatgtca	gcgatgccac	accgaatgta	gcagcttcgg	ctgcccata	gcaagactca	1860
cttgaggat	ttatttcaaa	tacggattca	agaatttcta	attatatgcc	aggtgactat	1920
caactgagat	ttaatcgtgc	cggcgtgcat	acatatgggt	cgactgggtc	gattgaaggt	1980
atgacctttg	cggatacaaga	tggtccaata	ggttatcaag	ttgaagtgcg	tattccctta	2040
gaaaaatgaag	tctatgttgg	cagaagactt	ggttttgact	tacaagtcaa	tgatgcatgg	2100
gaagtggcg	gtacttctgg	ccgacaagca	tttgctaaat	ggaatgatca	cactgacaat	2160
ggctggcagt	ctacagagtt	ttggggctgg	ttattattac	aaggcgatgc	ggcacctgtc	2220
ttaccggttg	tattagtggg	agaaggcttt	gaaacggatt	taggttcatt	ccaaccaagg	2280
ggtagtagta	cactgactcg	aaccacaagag	gttagtcatg	aaggcgacta	ttccgtattg	2340
gttagcaacc	gtgtcaacaa	ctggaacggg	gcgtcattac	cgtaacagg	cattgttcaa	2400
ccaggcaaca	cttatgagtt	tggtgggtac	attagagcaa	aagcagatgt	aactggatca	2460
tatatcatga	gtgggtgagtt	taataatgga	tctgggtgat	tagaaaacgg	tagtattaat	2520
cgatggccat	ggctatcaaa	ccgttcatta	acgatagcag	atggttttgt	tgagttaaag	2580
tcagaactaa	ccatacctag	tgacatgacg	acgtttaact	tgaactttga	acaccaaagt	2640
gctgaagtag	aattttactt	agatgctgtt	caagtcactt	tgattgcaga	agctgatgta	2700
acaccagtg	acccaccagt	agaccacca	gtagagccag	aaattacagt	ggtctattca	2760
atggtagatg	atgcagccat	tcaagggatt	gaagtgggaa	caacaggcac	tgctgaagat	2820
ttttcggata	ttagtgaagc	tttattagta	tctggttcac	cagttgttac	tgctgtagca	2880
catccagaag	aagcaggaaa	gatcgggtata	gagcttagta	atcgagcaga	gaattggcat	2940
gcgctagact	ttatgtttccc	agccataggt	gtgcagcggg	gtgggagcta	tcgatttgtg	3000
gccagtggcc	gtatggcaga	aggaacaggt	aattcaaadc	gtagaatgca	gtggaatcaa	3060
acggatgcgc	catggagtga	aatatcaggt	tctagaacca	atgtggcacc	tgacgcaacc	3120
acatggacca	ttgacgtgac	tttaagtcca	ttacagatca	acacattatt	aaacgctggg	3180
caaagaggtc	ttcgaattca	aacgggtaac	gcaccaactg	tgaccattac	cattgacgat	3240
gtgtttgttt	atcagattgg	tgacattgac	acagcaggtt	taccattacc	accacaatgg	3300
aattttgatt	tgccaagatt	atcagaatta	ttcgagccat	attttggctt	tggtaacatt	3360
tattcaaccg	aaacatttaa	gaacgctaatt	gaaacaaaaa	gagcattttt	acatcacttt	3420
aacgtgatta	cagcagaaaa	tggtcataaa	ccatccagta	ttgcagggcc	agaaaaatagc	3480
tttacagtac	cagaacctga	gcaattcaac	tttacggatg	cagaccgaat	tgtaaaccttt	3540
gctgttgaaa	atgacattga	attagtagga	catgcacttg	tatggcattc	acaaagtcca	3600
aactggctgt	ttagaagtgc	ggctaacaca	ccgctaacaa	gagcagaagc	caaagagcgc	3660
atggcatatt	acatgaaacg	tgtttcagag	cattttcgaag	cacaaggtac	attaggcgca	3720
ttttatgggt	gggatgttgt	gaatgaagcc	atcgccagtg	gtgggtgtac	attcgtagat	3780
caaccaggtc	attggcgcac	gcaaattgca	acatcatcac	catgggtcca	agcatttaac	3840
aatggattag	atgtagaagc	cgggtgaacat	gccagtgtat	atattttcta	tgcatactat	3900
tatgcaagaa	agtattttccc	aacatcgatc	ctatactaca	acgattacaa	cgatgaaata	3960
ccaaacaagc	gagacaatat	cgctcaaatg	gtagaagaga	taaatgcact	ttgggaagca	4020
catgaagaat	atgatggctg	cttactgatt	gaatccatcg	gtatgcaaag	tcattatcac	4080
atggaagggt	ggacaaccag	cgtagacaat	gtaagagctg	ctttagatcg	atacattgca	4140
acagggtgca	gagtcagttg	gactgagtta	gatatacatt	atgggtgtca	tggtagtaac	4200
gcataatgcat	cacttacacc	agaacaatta	gcggcacaa	cggagcgata	tgacagagata	4260
tttacattgt	atttagagcg	tgcatatcag	tttaagccgtg	tatccatctg	gggtatgtct	4320
gatgctaaca	gctggagaag	ttctggattc	ccattactat	ttgacagtcc	acttaattgct	4380
aaaccagcat	ttaatgccat	tgtagaatta	gttaaaaaact	gggagacacc	aacagttgta	4440
gcaccagtga	ttcaaacaag	aacactagca	ccattagaaa	gtgggtgaaag	agtcctttacc	4500
atgttagatg	tggttaagagg	atctaattgca	cctgtattgt	ttagcataac	agacgggtgca	4560
ttaccagaag	gtataatcct	tcattctaga	acaggtattt	tagaaggaac	accagttgaa	4620
gatggctact	atagctttac	tgtaactgct	agaaattacg	gcggttcaac	aagtcaagcg	4680
ctgactttaa	cagtaggtca	tccagtagca	ccaccagtaa	cgccaccagt	aacgccacca	4740
accgtaatac	ttgatgaatc	ggatatacca	caggctgtgc	caggccttag	ggcaccacag	4800
attgttgtaa	ccgttcaaga	aggcagtgaa	gtaacgtttg	atcttgaaaa	attagaagaa	4860
gttatggcat	cactttcaag	tcaagtgcc	ttgggtgttag	atgttgaaat	ggaagattct	4920
atcatcacct	tggtcaaac	attacttaaa	cgattaacag	acaaggcggc	tggaatcgaa	4980
atacaagcag	atggatttag	ttatatgctt	ccagcagagg	tattagaggc	aattctttgg	5040

<210> 362
 <211> 1680
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(26)

<400> 362
 Met Ala Arg Ser Lys Arg Val Leu Ala Trp Ile Met Ser Ser Val Leu
 1 5 10 15
 Leu Ile Ser Met Ala Met Pro Ser Phe Ala Ser Gly Asp Ser Ser Gln
 20 25 30
 Val Pro Arg Val Ile Phe Glu Thr Gly Phe Glu Thr Gly Leu Asp Gly
 35 40 45
 Phe Lys Gly Arg Gly Ser Ala Thr Leu Thr Arg Thr Thr Asp Glu Thr
 50 55 60
 Gln Ala Gly Asp Tyr Ser Val Leu Val Ser Asn Arg Leu Glu His Trp
 65 70 75 80
 Asn Gly Ala Ser Leu Pro Leu Thr Gly Phe Val Leu Pro Gly Asn Thr
 85 90 95
 Tyr Glu Phe Val Gly Tyr Ile Lys Ala Lys Ala Asp Val Ala Asp Asn
 100 105 110
 Tyr Val Met Ser Gly Glu Tyr Asn Glu Gly Ile Ser Gly Asn Gln Tyr
 115 120 125
 Pro Trp Ile Ser Asn Arg Leu Thr Val Gln Asp Gly Phe Val Glu
 130 135 140
 Phe Arg Gly Glu Leu Thr Ile Leu Glu Asp Met Thr Ser Phe Asn Leu
 145 150 155 160
 Asn Phe Glu His Gln Asn Ala Glu Val Glu Phe Tyr Leu Asp Ser Val
 165 170 175
 Gln Val Ile Leu Ile Glu Glu Gly Gln Val Asn Asp Leu Pro Met Asn
 180 185 190
 Val Arg Arg Ala Pro Leu Thr Leu Ala Glu Thr Pro Leu His Glu Ile
 195 200 205
 Trp Ala Asp His Phe Thr Ile Gly Asn Ile Tyr Thr Pro Gly Phe Arg
 210 215 220
 Thr Asp Ile Arg Gly Glu Val Leu Ala His His Phe Asn Val Ile Thr
 225 230 235 240
 Ala Glu Asn Ile Met Lys Pro Asp His Leu Gln Arg Glu Gln Gly Ile
 245 250 255
 Phe Thr Phe Ser Ala Ser Asn Asp Met Met Glu Phe Ala Arg Ala Asn
 260 265 270
 Asn Gln Glu Val Ile Gly His Thr Leu Val Trp His Ser Gln Ser Phe
 275 280 285
 Pro Trp Phe Glu Ala Leu Asn Pro Thr Arg Asp Glu Ala Ile Ala Ile
 290 295 300
 Met His Ala His Ile Glu Thr Val Met Gly His Phe Asn Glu Asn Tyr
 305 310 315 320
 Pro Gly Val Ile Thr Gly Trp Asp Val Leu Asn Glu Ala Ile Gln Pro
 325 330 335
 Arg Gln Gly Gln Asp Pro Glu Asn Trp Arg Leu His Leu Arg Asp Thr
 340 345 350
 Lys Trp Leu Arg Ala Ile Gly Asp Asp Tyr Ile Ala Ile Ala Phe Asn
 355 360 365
 Lys Ala His Glu Met Asp Pro Asp Ala Ile Leu Tyr Tyr Asn Asp Tyr
 370 375 380
 Asn Asp Asn Asp Tyr Phe Lys Ala Thr Ile Ile Lys Ala Met Val Gln
 385 390 395 400
 Glu Leu Arg Asn Glu Gly Val Pro Ile His Arg Ile Gly Met Gln Gly
 405 410 415
 His Tyr Asn Leu Gln Thr Pro Leu Asn Ser Ile Arg Thr Ser Val Glu
 420 425 430
 Arg Phe Ser Glu Ile Thr Gly His Glu Asp Leu Pro Pro Ile Gly Ile
 435 440 445
 Ser Phe Thr Glu Ile Asp Val Thr Val Pro Gly Phe Glu Ser Ala Ala
 450 455 460
 Arg Leu Pro Glu Glu Val Glu Ile Arg Gln Ala Gln Phe Tyr Ala Gln
 465 470 475 480
 Leu Met Gln Ile Leu Arg Asp Asn Ser Asp Val Ile His Arg Val Thr
 485 490 495
 Phe Trp Gly Met Ser Asp Arg Glu Ser Trp Arg Ala Asp Arg His Pro
 500 505 510
 Asn Met Leu Asp Pro Gln Tyr Gly Pro Lys His Val Phe His Ala Ile
 515 520 525
 Ala Asn Pro Glu Ala Phe Leu Thr Ala Tyr Pro Leu Pro Glu Thr Pro

530 535 540
 Asp Ala Gln Thr Ala Tyr Ala Ser Gln Gly Gln Pro Val Val Gly Gln
 545 550 555
 Phe Asn Leu Asp Ala Tyr Gln Asn Ser Glu Val Ile Pro Val Ala Asn
 565 570 575
 Gln Met Thr Ala His Asn Gly Ala Thr Ala Val Ala Arg Val Val Trp
 580 585 590
 His Glu Asp Ala Ile Tyr Ile Leu Ala Asn Val Ser Asp Ala Thr Pro
 595 600 605
 Asn Val Ala Ala Ser Ala Ala His Glu Gln Asp Ser Leu Glu Val Phe
 610 615 620
 Ile Ser Asn Thr Asp Ser Arg Ile Ser Asn Tyr Met Pro Gly Asp Tyr
 625 630 635
 Gln Leu Arg Phe Asn Arg Ala Gly Val His Thr Tyr Gly Ser Thr Gly
 645 650 655
 Ser Ile Glu Gly Met Thr Phe Ala Val Gln Asp Gly Pro Ile Gly Tyr
 660 665 670
 Gln Val Glu Val Arg Ile Pro Leu Glu Asn Glu Val Tyr Val Gly Arg
 675 680 685
 Arg Leu Gly Phe Asp Leu Gln Val Asn Asp Ala Trp Glu Val Gly Gly
 690 695 700
 Thr Ser Gly Arg Gln Ala Phe Ala Lys Trp Asn Asp His Thr Asp Asn
 705 710 715
 Gly Trp Gln Ser Thr Glu Phe Trp Gly Trp Leu Leu Gln Gly Asp
 725 730 735
 Ala Ala Pro Val Leu Pro Val Val Leu Val Glu Glu Gly Phe Glu Thr
 740 745 750
 Asp Leu Gly Ser Phe Gln Pro Arg Gly Ser Ser Thr Leu Thr Arg Thr
 755 760 765
 Gln Glu Val Ser His Glu Gly Asp Tyr Ser Val Leu Val Ser Asn Arg
 770 775 780
 Val Asn Asn Trp Asn Gly Ala Ser Leu Pro Leu Thr Gly Ile Val Gln
 785 790 795
 Pro Gly Asn Thr Tyr Glu Phe Val Gly Tyr Ile Arg Ala Lys Ala Asp
 805 810 815
 Val Thr Gly Ser Tyr Ile Met Ser Gly Glu Phe Asn Asn Gly Ser Gly
 820 825 830
 Val Leu Glu Asn Gly Ser Ile Asn Arg Trp Pro Trp Leu Ser Asn Arg
 835 840 845
 Ser Leu Thr Ile Ala Asp Gly Phe Val Glu Phe Lys Ser Glu Leu Thr
 850 855 860
 Ile Pro Ser Asp Met Thr Thr Phe Asn Leu Asn Phe Glu His Gln Asn
 865 870 875
 Ala Glu Val Glu Phe Tyr Leu Asp Ala Val Gln Val Thr Leu Ile Ala
 885 890 895
 Glu Ala Asp Val Thr Pro Val Asp Pro Val Asp Pro Pro Val Glu
 900 905 910
 Pro Glu Ile Thr Val Val Tyr Ser Met Val Asp Asp Ala Ala Ile Gln
 915 920 925
 Gly Ile Glu Val Gly Thr Thr Gly Thr Ala Glu Asp Phe Ser Asp Ile
 930 935 940
 Ser Glu Ala Leu Leu Val Ser Gly Ser Pro Val Val Thr Ala Val Ala
 945 950 955
 His Pro Glu Glu Ala Gly Lys Ile Gly Ile Glu Leu Ser Asn Arg Ala
 965 970 975
 Glu Asn Trp His Ala Leu Asp Phe Met Phe Pro Ala Ile Gly Val Gln
 980 985 990
 Arg Gly Gly Ser Tyr Arg Phe Val Ala Ser Gly Arg Met Ala Glu Gly
 995 1000 1005
 Thr Gly Asn Ser Asn Arg Arg Met Gln Trp Asn Gln Thr Asp Ala Pro
 1010 1015 1020
 Trp Ser Glu Ile Ser Gly Ser Arg Thr Asn Val Ala Pro Ala Ala Thr
 1025 1030 1035
 Thr Trp Thr Ile Asp Val Thr Leu Ser Arg Leu Gln Ile Asn Thr Leu
 1045 1050 1055
 Leu Asn Ala Gly Gln Arg Gly Leu Arg Ile Gln Thr Gly Asn Ala Pro
 1060 1065 1070
 Thr Val Thr Ile Thr Ile Asp Asp Val Phe Val Tyr Gln Ile Gly Asp
 1075 1080 1085

Ile Asp Thr Ala Gly Leu Pro Leu Pro Pro Gln Trp Asn Phe Asp Leu
 1090 1095 1100
 Pro Arg Leu Ser Glu Leu Phe Glu Pro Tyr Phe Gly Leu Gly Asn Ile
 1105 1110 1115 1120
 Tyr Ser Thr Glu Thr Leu Met Asn Ala Asn Glu Thr Lys Arg Ala Phe
 1125 1130 1135
 Leu His His Phe Asn Val Ile Thr Ala Glu Asn Gly His Lys Pro Ser
 1140 1145 1150
 Ser Ile Ala Gly Pro Glu Asn Ser Phe Thr Val Pro Glu Pro Glu Gln
 1155 1160 1165
 Phe Asn Phe Thr Asp Ala Asp Arg Ile Val Asn Phe Ala Val Glu Asn
 1170 1175 1180
 Asp Ile Glu Leu Val Gly His Ala Leu Val Trp His Ser Gln Ser Pro
 1185 1190 1195 1200
 Asn Trp Leu Phe Arg Ser Ala Ala Asn Thr Pro Leu Thr Arg Ala Glu
 1205 1210 1215
 Ala Lys Glu Arg Met Ala Tyr Tyr Met Lys Thr Val Ser Glu His Phe
 1220 1225 1230
 Glu Ala Gln Gly Thr Leu Gly Ala Phe Tyr Gly Trp Asp Val Val Asn
 1235 1240 1245
 Glu Ala Ile Ala Ser Gly Gly Gly Thr Phe Val Asp Gln Pro Gly His
 1250 1255 1260
 Trp Arg Thr Gln Met Arg Thr Ser Ser Pro Trp Phe Gln Ala Phe Asn
 1265 1270 1275 1280
 Asn Gly Leu Asp Val Glu Ala Gly Glu His Ala Ser Asp Tyr Ile Phe
 1285 1290 1295
 Tyr Ala Tyr Tyr Tyr Ala Arg Lys Tyr Phe Pro Thr Ser Ile Leu Tyr
 1300 1305 1310
 Tyr Asn Asp Tyr Asn Asp Glu Ile Pro Asn Lys Arg Asp Asn Ile Ala
 1315 1320 1325
 Gln Met Val Glu Glu Ile Asn Ala Leu Trp Glu Ala His Glu Glu Tyr
 1330 1335 1340
 Asp Gly Arg Leu Leu Ile Glu Ser Ile Gly Met Gln Ser His Tyr His
 1345 1350 1355 1360
 Met Glu Gly Trp Thr Thr Ser Val Asp Asn Val Arg Ala Ala Leu Asp
 1365 1370 1375
 Arg Tyr Ile Ala Thr Gly Ala Arg Val Ser Val Thr Glu Leu Asp Ile
 1380 1385 1390
 Thr Tyr Gly Gly His Gly Ser Asn Ala Tyr Ala Ser Leu Thr Pro Glu
 1395 1400 1405
 Gln Leu Ala Ala Gln Ala Glu Arg Tyr Ala Glu Ile Phe Thr Leu Tyr
 1410 1415 1420
 Leu Glu Arg Ala Asp Gln Leu Ser Arg Val Ser Ile Trp Gly Met Ser
 1425 1430 1435 1440
 Asp Ala Asn Ser Trp Arg Ser Ser Gly Phe Pro Leu Leu Phe Asp Ser
 1445 1450 1455
 Ser Leu Asn Ala Lys Pro Ala Phe Asn Ala Ile Val Glu Leu Val Lys
 1460 1465 1470
 Asn Trp Glu Thr Pro Thr Val Val Ala Pro Val Ile Gln Thr Arg Thr
 1475 1480 1485
 Leu Ala Pro Leu Glu Ser Gly Glu Arg Val Phe Thr Met Leu Asp Val
 1490 1495 1500
 Val Arg Gly Ser Asn Ala Pro Val Trp Phe Ser Ile Thr Asp Gly Ala
 1505 1510 1515 1520
 Leu Pro Glu Gly Ile Ile Leu His Ser Arg Thr Gly Ile Leu Glu Gly
 1525 1530 1535
 Thr Pro Val Glu Asp Gly His Tyr Ser Phe Thr Val Thr Ala Arg Asn
 1540 1545 1550
 Tyr Gly Gly Ser Thr Ser Gln Ala Leu Thr Leu Thr Val Gly His Pro
 1555 1560 1565
 Val Ala Pro Pro Val Thr Pro Pro Val Thr Pro Pro Thr Val Ile Ile
 1570 1575 1580
 Asp Glu Ser Asp Ile Pro Gln Ala Gly Pro Gly Leu Arg Ala Pro Gln
 1585 1590 1595 1600
 Ile Val Val Thr Val Gln Glu Gly Ser Glu Val Thr Phe Asp Leu Glu
 1605 1610 1615
 Lys Leu Glu Glu Val Met Ala Ser Leu Ser Ser Gln Val Pro Leu Val
 1620 1625 1630
 Leu Asp Val Glu Leu Glu Asp Ser Ile Ile Thr Leu Asp Gln Thr Leu

1635 1640 1645
 Leu Lys Arg Leu Thr Asp Lys Ala Ala Gly Ile Glu Ile Gln Ala Asp
 1650 1655 1660
 Gly Phe Ser Tyr Met Leu Pro Ala Glu Val Leu Glu Ala Ile Leu Trp
 1665 1670 1675 1680

<210> 363
 <211> 1317
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 363
 gtgaccaccc gcgcacaggt tcttgatcag gcgctggcac tgggccaccg gtgtggctgg 60
 gaaaaactca gcctggacgc catcgccagg gcgctgggcc gttttggcag cctggctctg 120
 tgtgtcgcgc tgttgagcgc ctgcggcagt agtagtagct ccctggatga tccgggtgct 180
 ggcagcagtt cttccagctc tgagagcagc caaagctcca gcgccagttc ccaggctgat 240
 ggcgacggta cccaggacag cctctacgcc caggcggact tccctgtagg gggtgcggtg 300
 caggtggcca attgggagcc tttcagcctg tttaccgcgc ccgatgccgc tgcgcgtcag 360
 aacctggttg cccgacactt ctccgaagtg accgcgacca acgtcatgaa aatgtcctat 420
 atgcgcacca acagtgggtg ttttaccgac gcgcccggcg gtccgctgat tgattttgcc 480
 gcgcgccaatg gcatcaaagt gcacggtcac gcaactggtc ggcatgcgga ttatcagggtg 540
 ccaaattgtgt ttcgtgacta cgaaggggac aattggcagg ggcttttaac cgagcatgtc 600
 gagggcggtta tggggctggt tgacgacacc gtggttaagt gggatgtcgt aaacgaagcg 660
 gttgataccg gctcacctga cggctggcgc cggctcattt tctataattt tgcgcccgcg 720
 gaagcagggc aggtgccgga atatattgaa gtggcttacc aggccgctcg agaggccaat 780
 ccggaagtga cctctacta caacgatttt gacaacacgg ccaataccgg gcgcctcaac 840
 aagaccctgg aaattgccga tcgcctgaaa gagctggacg cgatcgacgg tatcgggttc 900
 cagatgcacg cctatatgaa ctacccgagt attgcgcagt ttcgcaatgc ctttcaggaa 960
 gtggtcgatc gtgacctgaa agtcaaagtc accgagctgg acattgccat cgtcaaccct 1020
 tacggcagct cgacgcctcc gccgctgccg gagtttgatc aggcgctggc cgacgcccaa 1080
 ggtgtccgtt actgccagat tgccgagggc tatctggatg tcgttcctgc cgagctgcgg 1140
 ggtggtttca ccgtctgggg cctgaccgat gacgacagct ggctgatggg agcgttcgcg 1200
 tccgcaaccg gcgccaataa cgaccagggtc tatccggtgt tgtttgacga taatctgcaa 1260
 gccaaagcccg cgttcttttg cgtaacgcgc gccctccgcg gcgaaccctg cgagtaa 1317

<210> 364
 <211> 438
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 364
 Met Thr Thr Arg Ala Gln Val Leu Asp Gln Ala Leu Ala Leu Gly His
 1 5 10 15
 Arg Cys Gly Trp Glu Lys Leu Ser Leu Asp Ala Ile Ala Arg Ala Leu
 20 25 30
 Gly Arg Phe Gly Ser Leu Ala Leu Cys Val Ala Leu Leu Ser Ala Cys
 35 40 45
 Gly Ser Ser Ser Ser Leu Asp Asp Pro Gly Ala Gly Ser Ser Ser
 50 55 60
 Ser Ser Ser Glu Ser Ser Gln Ser Ser Ser Ala Ser Ser Gln Ala Asp
 65 70 75 80
 Gly Asp Gly Thr Gln Asp Ser Leu Tyr Ala Gln Ala Asp Phe Pro Val
 85 90 95
 Gly Val Ala Val Gln Val Ala Asn Trp Glu Pro Phe Ser Leu Phe Thr
 100 105 110
 Ala Pro Asp Ala Ala Ala Arg Gln Asn Leu Val Ala Arg His Phe Ser
 115 120 125
 Glu Val Thr Ala Thr Asn Val Met Lys Met Ser Tyr Met Arg Thr Asn
 130 135 140
 Ser Gly Gly Phe Thr Asp Ala Pro Ala Arg Pro Leu Ile Asp Phe Ala
 145 150 155 160
 Arg Ala Asn Gly Ile Lys Val His Gly His Ala Leu Val Trp His Ala
 165 170 175

Asp Tyr Gln Val Pro Asn Val Phe Arg Asp Tyr Glu Gly Asp Asn Trp
 180 185 190
 Gln Gly Leu Leu Thr Glu His Val Glu Gly Val Met Gly Leu Phe Asp
 195 200 205
 Asp Thr Val Val Ser Trp Asp Val Val Asn Glu Ala Val Asp Thr Gly
 210 215 220
 Ser Pro Asp Gly Trp Arg Arg Ser Ile Phe Tyr Asn Phe Ala Pro Pro
 225 230 235 240
 Glu Ala Gly Gln Val Pro Glu Tyr Ile Glu Val Ala Tyr Gln Ala Ala
 245 250 255
 Arg Glu Ala Asn Pro Glu Val Thr Leu Tyr Tyr Asn Asp Phe Asp Asn
 260 265 270
 Thr Ala Asn Thr Gly Arg Leu Asn Lys Thr Leu Glu Ile Ala Asp Arg
 275 280 285
 Leu Lys Glu Leu Asp Ala Ile Asp Gly Ile Gly Phe Gln Met His Ala
 290 295 300
 Tyr Met Asn Tyr Pro Ser Ile Ala Gln Phe Arg Asn Ala Phe Gln Glu
 305 310 315 320
 Val Val Asp Arg Asp Leu Lys Val Lys Val Thr Glu Leu Asp Ile Ala
 325 330 335
 Ile Val Asn Pro Tyr Gly Ser Ser Thr Pro Pro Pro Leu Pro Glu Phe
 340 345 350
 Asp Gln Ala Leu Ala Asp Ala Gln Gly Val Arg Tyr Cys Gln Ile Ala
 355 360 365
 Glu Ala Tyr Leu Asp Val Val Pro Ala Glu Leu Arg Gly Gly Phe Thr
 370 375 380
 Val Trp Gly Leu Thr Asp Asp Asp Ser Trp Leu Met Gly Ala Phe Ala
 385 390 395 400
 Ser Ala Thr Gly Ala Gln Tyr Asp Gln Val Tyr Pro Val Leu Phe Asp
 405 410 415
 Asp Asn Leu Gln Ala Lys Pro Ala Phe Phe Gly Val Lys Arg Ala Leu
 420 425 430
 Arg Gly Glu Pro Cys Glu
 435

<210> 365
 <211> 3246
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 365
 atgaaccact tgcgttcaaa atcgctgcgc atggcgtggc aaccgggact gcttgcgaca 60
 accgtgtctgc cgttggcggc tgccgcccc ataccagcgc cgaatacggg taccaaagtg 120
 agcaatactt cgtccatcac cactccggct gcggctccac agtcgcagcc acaaccaacg 180
 caagacgcaa acgctccgc accgcttaaa gcggctttcc gggataagt tctcatcggc 240
 gcggtgtga gtgacgtgc gctgcgaggc agtgcgccg acaaggtggc gatagccacc 300
 acgcacttta acgcgtcac cgccgaaaac gccatgaagc cagacgcgat gcaaccgcgc 360
 gaagggcagt tcaacttcgc tgcaggcgat cggctcgtcg aactcgccga aaaaagcggc 420
 ggtgtgccc tgcgccacac gctgggtgtg cagcgcaaa caccgaagt gtttttgaa 480
 gggccggatg gacagccgc gacgcgcgaa ctggctttgg agcgcgtgc caaacacatt 540
 tccactgtgg tggggcgcta caaaggcgc atcaaggagt gggatgtgg gaacgaagcc 600
 atcaacgacg gaccgggtgt gctgcgtccc tctcccggc tcaaagccat cggcgaagat 660
 tacatcgcc aagccttcg cgccgcgcac gccgcgacc ccgacgcgat ttgatttat 720
 aacgattaca acatcgaact gggctacaaa cgctgcaact cctaaaatcg 780
 ctcatgacc agaaagtgc gattcacgcc gtgggcattc agggctactg gcgcatggac 840
 aaccggaact tgcgcgaagt ggaacaggcc atcaaagagt ttccggcgt ggggttgaaa 900
 gtcattgatca ccgaactcga catcggcgtg ctgccgacgc gttatcaggg gcgcatatt 960
 tcagcgacc aaacctgac gccgaacag cgcccggtga tgaacccta tacggacgga 1020
 ttgcccggac atgtggcgca aaagcacgcc gagcgtatc gccaggcgtt tgagatgttc 1080
 ctgcccacaa aagacaaaat cagtcgtgtg acattttggg gtgtggacga cggcacttcg 1140
 tggctgaacg gtttcccggt gcgcggccgc accgattatc cgctgctatt tgatcgctcag 1200
 ggcaagccaa aacccgcctt ttgcgcgtg caaacgcgg cgatgggcgc aacagcgcaa 1260
 ccgagcgcca gcgtcccgcc aacgcattgc gccgctctg catccacaa cattcgcggc 1320
 gccgagtttc ctgcgtgga aagcgacggg cgggtgacgt ttcgcatcaa agcgcctgac 1380
 gcgcaaaaag tgcaatttga tttaggtaag cttacgacg ccaccgcga cgcgagggc 1440
 aactggacgg cgaccacaga gccacaagt cccggttcc attattattt ttgattgtc 1500

```

gatggagtgc gcgtggccga cccggcgagc gaaacctttt acggtgcggg ccgccagatg 1560
agcggcatcg aaattcccga tcccgcagac gcgtttttatt cgccgcaaaa cgtgccgcat 1620
ggcgaagtgc gcgaacgctg gtatittttcc aacaccacgc aggcgtggcg gcgcatcttc 1680
atttatacgc cgccgggtta cgacaccgat caggccatgc gttttcctgt gctgtatttg 1740
cagcacggcg gtggcgaaga cgaacgcggc tggcccaatc aggggcgcg gagctttatc 1800
atggacaatc tcatcgcgca gggcaaaagcc aaaccgatgc tgggtggtgat ggagcaaggc 1860
tatgcgcgca agcccgatga accgcagggtg ccgctgcgcc cgcccggaa gcaacgccga 1920
gcgatgccgc ccgactttta tcgcatgttc gccacgctgg gcgaagtgtt caccaaagac 1980
ctgattccgt ttattgacgc aaattaccgc accaaaaccg agcgcgaaaa ccgcgcgatg 2040
gccggacttt cgatgggtgg aatgcaaagt ttcatcatcg gcctggcgaa caccgatcta 2100
ttcgcgcacc tcggcggttt cagcggcgcg ggtggtggtt ttggcgggcg cgctttcgac 2160
gccaaaaccg cgcacggcgg tgtgatggcc gatgccgatg ctttcaacaa aaaagttcgc 2220
acgatgtttc tcagcatcgg cactgccgag aacgagcgtt ttcagagcag cgtgcgcggg 2280
taccgcgacg cgtgaccaa agcgggcatc aaaaccacgt tctacgaatc gcccggcact 2340
tcgcacgagt ggctgacatg gcgcccgcag ctgcgcgaat tcgcgcgcgt cttgtttcaa 2400
gaggccaaca cgcagatcga gcgcggcccc aatgcccgcc cgattgcgcc gcagccgatt 2460
gttcttggcc cgggcgacaa gccgccttc cctccggcgc cctccggttt cgatgcgcgg 2520
cgcatggca ttccgcacgg cgaaattaa cttgtggaat acccttctgc cacggtcggc 2580
accacgcgca agatgcaggt ctatacgccg ccgggctaca acccgcaaga agaatatccc 2640
gtgctctatt tgctgcacgg catcggcggg gacgagtggg aatggaaaaa tggcggcacg 2700
cccgaagtga ttctcgacaa cctctacgct gagaagaaac tccagccgat gatcgtggtg 2760
atgcccaatg ggcgcgcgca aaaagacgac cgtcctatcg gcaacgtgtt cgcttccgt 2820
ccggcggttg cgacgtttga gaaagatttg ctgaacgaca ttatccccct tgttgagaag 2880
aattaccga ccaaaaccgg cccgcaaaat cgcgctttgg ccggtctttc gatggcgggc 2940
gggcaatctc tcaacttgg cctcggcaac ctgcacact tcgctgggtt tggcggttt 3000
tcgtccgcgc ccaacacgcg cagcggcgca agtctactgg ccaatcccga cgacgcaaa 3060
aagaagctga agctgctgtg ggtttcgtgc ggcgataaag acaatttgat gtttatcagc 3120
cagcgcacgc accgttatct tgccgagaat aacgtgccgc acatctggca tgtacagccc 3180
ggcggaacag acttcaagggt gtggaagcaa gacctgtata acttcgcca actgctattc 3240
cgtaa

```

<210> 366

<211> 1081

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<221> SIGNAL

<222> (1)...(65)

<400> 366

```

Met Asn His Phe Ala Ser Lys Ser Leu Arg Met Ala Trp Gln Pro Gly
1      5      10      15
Leu Leu Ala Thr Val Leu Pro Leu Ala Ala Ala Ala Pro Ile Pro
20      25      30
Ala Pro Asn Thr Asp Thr Lys Val Ser Asn Thr Ser Ser Ile Thr Thr
35      40      45
Pro Ala Ala Ala Pro Gln Ser Gln Pro Gln Pro Thr Gln Asp Ala Asn
50      55      60
Ala Pro Ala Pro Leu Lys Ala Ala Phe Arg Asp Lys Phe Leu Ile Gly
65      70      75      80
Ala Val Leu Ser Asp Ala Ala Leu Arg Gly Ser Ala Pro Asp Lys Val
85      90      95
Ala Ile Ala Thr His Phe Asn Ala Leu Thr Ala Glu Asn Ala Met
100     105     110
Lys Pro Asp Ala Met Gln Pro Arg Glu Gly Gln Phe Asn Phe Ala Ala
115     120     125
Gly Asp Arg Leu Val Glu Leu Ala Glu Lys Ser Gly Gly Val Pro Ile
130     135     140
Gly His Thr Leu Val Trp His Ala Gln Thr Pro Lys Trp Phe Phe Glu
145     150     155     160
Gly Pro Asp Gly Gln Pro Ala Thr Arg Glu Leu Ala Leu Glu Arg Met
165     170     175
Arg Lys His Ile Ser Thr Val Val Gly Arg Tyr Lys Gly Arg Ile Lys
180     185     190
Glu Trp Asp Val Val Asn Glu Ala Ile Asn Asp Gly Pro Gly Val Leu
195     200     205

```

Arg	Pro	Ser	Pro	Trp	Leu	Lys	Ala	Ile	Gly	Glu	Asp	Tyr	Ile	Ala	Glu
	210					215					220				
Ala	Phe	Arg	Ala	Ala	His	Ala	Ala	Asp	Pro	Asp	Ala	Ile	Leu	Ile	Tyr
225					230					235					240
Asn	Asp	Tyr	Asn	Ile	Glu	Leu	Gly	Tyr	Lys	Arg	Pro	Lys	Ala	Leu	Gln
				245					250					255	
Leu	Leu	Lys	Ser	Leu	Ile	Asp	Gln	Lys	Val	Pro	Ile	His	Ala	Val	Gly
			260					265					270		
Ile	Gln	Gly	His	Trp	Arg	Met	Asp	Asn	Pro	Asn	Phe	Ala	Glu	Val	Glu
		275					280					285			
Gln	Ala	Ile	Lys	Glu	Phe	Ser	Ala	Leu	Gly	Leu	Lys	Val	Met	Ile	Thr
	290					295					300				
Glu	Leu	Asp	Ile	Gly	Val	Leu	Pro	Thr	Arg	Tyr	Gln	Gly	Ala	Asp	Ile
305					310					315					320
Ser	Ala	Thr	Glu	Thr	Met	Thr	Pro	Glu	Gln	Arg	Ala	Val	Met	Asn	Pro
				325					330					335	
Tyr	Thr	Asp	Gly	Leu	Pro	Asp	Asp	Val	Ala	Gln	Lys	His	Ala	Glu	Arg
			340					345					350		
Tyr	Arg	Gln	Ala	Phe	Glu	Met	Phe	Leu	Arg	His	Lys	Asp	Lys	Ile	Ser
		355					360					365			
Arg	Val	Thr	Phe	Trp	Gly	Val	Asp	Asp	Gly	Thr	Ser	Trp	Leu	Asn	Gly
	370					375					380				
Phe	Pro	Val	Arg	Gly	Arg	Thr	Asp	Tyr	Pro	Leu	Leu	Phe	Asp	Arg	Gln
385					390					395					400
Gly	Lys	Pro	Lys	Pro	Ala	Phe	Phe	Ala	Val	Gln	Asn	Ala	Ala	Met	Gly
				405					410					415	
Ala	Thr	Ala	Gln	Pro	Ser	Ala	Ser	Ala	Pro	Ala	Thr	His	Gly	Ala	Ala
			420					425					430		
Pro	Ala	Ser	Thr	Asn	Ile	Arg	Gly	Ala	Glu	Phe	Pro	Arg	Val	Glu	Ser
		435					440					445			
Asp	Gly	Arg	Val	Thr	Phe	Arg	Ile	Lys	Ala	Pro	Asp	Ala	Gln	Lys	Val
	450					455					460				
Gln	Phe	Asp	Leu	Gly	Lys	Pro	Tyr	Asp	Ala	Thr	Arg	Asp	Ala	Glu	Gly
465					470					475					480
Asn	Trp	Thr	Ala	Thr	Thr	Glu	Pro	Gln	Val	Pro	Gly	Phe	His	Tyr	Tyr
				485					490					495	
Phe	Leu	Ile	Val	Asp	Gly	Val	Arg	Val	Ala	Asp	Pro	Ala	Ser	Glu	Thr
			500					505					510		
Phe	Tyr	Gly	Ala	Gly	Arg	Gln	Met	Ser	Gly	Ile	Glu	Ile	Pro	Asp	Pro
		515					520					525			
Asp	Ser	Ala	Phe	Tyr	Ser	Pro	Gln	Asn	Val	Pro	His	Gly	Glu	Val	Arg
	530					535					540				
Glu	Arg	Trp	Tyr	Phe	Ser	Asn	Thr	Thr	Gln	Ala	Trp	Arg	Arg	Ile	Phe
545					550					555					560
Ile	Tyr	Thr	Pro	Pro	Gly	Tyr	Asp	Thr	Asp	Gln	Ala	Met	Arg	Phe	Pro
				565					570					575	
Val	Leu	Tyr	Leu	Gln	His	Gly	Gly	Gly	Glu	Asp	Glu	Arg	Gly	Trp	Pro
			580					585					590		
Asn	Gln	Gly	Arg	Val	Ser	Phe	Ile	Met	Asp	Asn	Leu	Ile	Ala	Gln	Gly
		595					600					605			
Lys	Ala	Lys	Pro	Met	Leu	Val	Val	Met	Glu	Gln	Gly	Tyr	Ala	Arg	Lys
	610					615					620				
Pro	Asp	Glu	Pro	Gln	Val	Pro	Leu	Arg	Pro	Pro	Gly	Ser	Asn	Ala	Gly
625					630					635					640
Ala	Met	Pro	Pro	Asp	Phe	Asn	Arg	Met	Phe	Ala	Thr	Leu	Gly	Glu	Val
				645					650					655	
Phe	Thr	Lys	Asp	Leu	Ile	Pro	Phe	Ile	Asp	Ala	Asn	Tyr	Arg	Thr	Lys
			660					665					670		
Thr	Glu	Arg	Glu	Asn	Arg	Ala	Met	Ala	Gly	Leu	Ser	Met	Gly	Gly	Met
		675					680					685			
Gln	Ser	Phe	Ile	Ile	Gly	Leu	Ala	Asn	Thr	Asp	Leu	Phe	Ala	His	Leu
	690					695					700				
Gly	Gly	Phe	Ser	Gly	Ala	Gly	Gly	Gly	Phe	Gly	Gly	Gly	Ala	Phe	Asp
705					710					715					720
Ala	Lys	Thr	Ala	His	Gly	Gly	Val	Met	Ala	Asp	Ala	Asp	Ala	Phe	Asn
				725					730					735	
Lys	Lys	Val	Arg	Thr	Met	Phe	Leu	Ser	Ile	Gly	Thr	Ala	Glu	Asn	Glu
			740					745					750		
Arg	Phe	Gln	Ser	Ser	Val	Arg	Gly	Tyr	Arg	Asp	Ala	Leu	Thr	Lys	Ala

755 760 765
 Gly Ile Lys Thr Thr Phe Tyr Glu Ser Pro Gly Thr Ser His Glu Trp
 770 775 780
 Leu Thr Trp Arg Arg Ser Leu Arg Glu Phe Ala Pro Leu Leu Phe Gln
 785 790 795 800
 Glu Ala Asn Thr Gln Ile Glu Arg Gly Pro Asn Ala Arg Pro Ile Ala
 805 810 815
 Pro Gln Pro Ile Val Leu Gly Pro Gly Asp Lys Pro Ala Phe Pro Pro
 820 825 830
 Ala Pro Ser Gly Phe Asp Ala Arg Asp Gly Ile Pro His Gly Glu
 835 840 845
 Ile Lys Leu Val Glu Tyr Pro Ser Ala Thr Val Gly Thr Thr Arg Lys
 850 855 860
 Met Gln Val Tyr Thr Pro Gly Tyr Asn Pro Gln Glu Glu Tyr Pro
 865 870 875 880
 Val Leu Tyr Leu Leu His Gly Ile Gly Gly Asp Glu Trp Glu Trp Lys
 885 890 895
 Asn Gly Gly Thr Pro Glu Val Ile Leu Asp Asn Leu Tyr Ala Glu Lys
 900 905 910
 Lys Leu Gln Pro Met Ile Val Val Met Pro Asn Gly Arg Ala Gln Lys
 915 920 925
 Asp Asp Arg Pro Ile Gly Asn Val Phe Ala Ser Ala Pro Ala Phe Ala
 930 935 940
 Thr Phe Glu Lys Asp Leu Leu Asn Asp Ile Ile Pro Phe Val Glu Lys
 945 950 955 960
 Asn Tyr Pro Thr Lys Thr Gly Pro Gln Asn Arg Ala Leu Ala Gly Leu
 965 970 975
 Ser Met Gly Gly Gln Ser Leu Asn Phe Gly Leu Gly Asn Leu Asp
 980 985 990
 Thr Phe Ala Trp Val Gly Gly Phe Ser Ser Ala Pro Asn Thr Arg Ser
 995 1000 1005
 Gly Ala Ser Leu Leu Ala Asn Pro Asp Asp Ala Lys Lys Lys Leu Lys
 1010 1015 1020
 Leu Leu Trp Val Ser Cys Gly Asp Lys Asp Asn Leu Met Phe Ile Ser
 1025 1030 1035 1040
 Gln Arg Thr His Arg Tyr Leu Ala Glu Asn Asn Val Pro His Ile Trp
 1045 1050 1055
 His Val Gln Pro Gly Gly His Asp Phe Lys Val Trp Lys Gln Asp Leu
 1060 1065 1070
 Tyr Asn Phe Ala Gln Leu Leu Phe Arg
 1075 1080

<210> 367
 <211> 1338
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 367
 atgaaaagaa ttggattact atttatggcg ttggcgctaa ccgcatttat ggcgcagcat 60
 tcgtccgctc aaaggatttg caataaccaa acagggaccc atgggtggatt ctactacaca 120
 tgggtggagt atgggggttg atctgcatgt ataacaatgg gcgatggcgg taactacagc 180
 acccaatgga gcaataaccg taactttgta ggcggttaagg gttggagcac aggaagatcc 240
 aaccgcgtaa ttagttacaa tgctggtaac tggctgccat cgggtaatgc ttacctatgt 300
 ttatatggct ggactaccaa cccgcttgtt gagtactacg tagttgatag ctggggttct 360
 tggagacctc ccggagcaac atcgcaggga acagtaaata ctgatggtgg cacctatgag 420
 atatacagaa ctcagcgtgt aaaccagcca tctattcagg ggaatactac tttctatcag 480
 tattggagcg ttagaacctc taaaaggggc actggaagca atgctaccat caccttcag 540
 aaccacgtaa atgcttgggc aagtaggggt tggaaacttg gagctcatag ctatcaggta 600
 ctggctaccg aggggttatca gagcagcggg agttcaaata ttactgtttg ggaagggtgg 660
 tcaagtggag gttcttcagg tgggaagcacc ggaggcagca ctggaggtgg atcacacgag 720
 atcattgtaa gagcccgttg ttagtagtagg tcagagcaaa ttaggcttag ggttggcaat 780
 acaaccgttg caacttggac cttactacc gttataggc actataggc tactacctca 840
 gctactgggt gtattctggg agagtacttc aatgataggc gcaaccgtga tgttcagatt 900
 gattacatta gggtaaacgg ctcaactcgt caatctgaga acatgtcgt caatacaggg 960
 gtatggcaga atggctcatg cggcggctcc aatagcgagt ggctacactg caacggagct 1020
 attggctacg gcgatgtggg tactggcaga tcaaccgctg ttgaggaagc atttactgct 1080

```

gccgaggatt gtggctgtga acctaaggca accctattcc ccaaccctgc tggcagtacc 1140
ctcagtatta tgctagacag gcaaccctat ggcgatgtaa gtattagaat atataatacg 1200
gtagggtcag ttgttcgcac catcaacaat ccagacctac tctctgaggt tgatgtcagt 1260
gcattaaatt ctggaatcta cttttagtag cttaggtccg aaggacatgt aagcaactac 1320
aaatttatta aaaagtag                                     1338

```

<210> 368
 <211> 445
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(23)

```

<400> 368
Met Lys Arg Ile Gly Leu Leu Phe Met Ala Leu Ala Leu Thr Ala Phe
1      5      10      15
Met Ala Gln His Ser Ser Ala Gln Arg Ile Cys Asn Asn Gln Thr Gly
20      25      30
Thr His Gly Phe Tyr Tyr Thr Trp Trp Ser Asp Gly Gly Ser
35      40      45
Ala Cys Ile Thr Met Gly Asp Gly Gly Asn Tyr Ser Thr Gln Trp Ser
50      55      60
Asn Thr Gly Asn Phe Val Gly Gly Lys Gly Trp Ser Thr Gly Arg Ser
65      70      75      80
Asn Arg Val Ile Ser Tyr Asn Ala Gly Asn Trp Ser Pro Ser Gly Asn
85      90      95
Ala Tyr Leu Cys Leu Tyr Gly Trp Thr Thr Asn Pro Leu Val Glu Tyr
100     105     110
Tyr Val Val Asp Ser Trp Gly Ser Trp Arg Pro Pro Gly Ala Thr Ser
115     120     125
Gln Gly Thr Val Asn Thr Asp Gly Gly Thr Tyr Glu Ile Tyr Arg Thr
130     135     140
Gln Arg Val Asn Gln Pro Ser Ile Gln Gly Asn Thr Thr Phe Tyr Gln
145     150     155     160
Tyr Trp Ser Val Arg Thr Ser Lys Arg Ala Thr Gly Ser Asn Ala Thr
165     170     175
Ile Thr Phe Gln Asn His Val Asn Ala Trp Ala Ser Arg Gly Trp Asn
180     185     190
Leu Gly Ala His Ser Tyr Gln Val Leu Ala Thr Glu Gly Tyr Gln Ser
195     200     205
Ser Gly Ser Ser Asn Ile Thr Val Trp Glu Gly Gly Ser Ser Gly Gly
210     215     220
Ser Ser Gly Gly Ser Thr Gly Gly Ser Thr Gly Gly Ser His Glu
225     230     235     240
Ile Ile Val Arg Ala Arg Gly Val Val Gly Ser Glu Gln Ile Arg Leu
245     250     255
Arg Val Gly Asn Thr Thr Val Ala Thr Trp Thr Leu Thr Thr Gly Tyr
260     265     270
Arg Asp Tyr Arg Ala Thr Thr Ser Ala Thr Gly Gly Ile Leu Val Glu
275     280     285
Tyr Phe Asn Asp Ser Gly Asn Arg Asp Val Gln Ile Asp Tyr Ile Arg
290     295     300
Val Asn Gly Ser Thr Arg Gln Ser Glu Asn Met Ser Tyr Asn Thr Gly
305     310     315     320
Val Trp Gln Asn Gly Ser Cys Gly Gly Ser Asn Ser Glu Trp Leu His
325     330     335
Cys Asn Gly Ala Ile Gly Tyr Gly Asp Val Val Thr Gly Arg Ser Thr
340     345     350
Ala Val Glu Glu Ala Phe Thr Ala Ala Glu Asp Cys Gly Cys Glu Pro
355     360     365
Lys Ala Thr Leu Phe Pro Asn Pro Ala Gly Ser Thr Leu Ser Ile Met
370     375     380
Leu Asp Arg Gln Pro Tyr Gly Asp Val Ser Ile Arg Ile Tyr Asn Thr
385     390     395     400
Val Gly Ala Val Val Arg Thr Ile Asn Asn Pro Asp Leu Leu Thr Glu

```


Val Asp Val Ser Ala Leu Asn Ser Gly Ile Tyr Phe Val Glu Leu Arg
 405 410 415
 Ser Glu Gly His Val Ser Asn Tyr Lys Phe Ile Lys Lys
 420 425 430 435 440 445

<210> 369
 <211> 1077
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 369
 atgaaatcat ttatcactgg caaaaaaatt gctgctggac taattactgc agctgctttg 60
 agcgcaccta tggtagcgc gcaaaccctg acttcaaatt ctcaaggcac ccacgacgga 120
 tttttctact ctttctggaa ggactcaggc aacgcctcaa tgaacttatt ggcgggcggc 180
 cgttatcagt ctactggaa caccggcacc aacaactggg taggcggtaa aggctggaac 240
 ccaggcacta acaaccgtgt aattaactac tctggttact acggtgtgga caactcccaa 300
 aactcttacg tcgcgcttta cggctggacc agaaacccat tgggtgagta ctactgtatt 360
 gagagctacg gctcatataa ccctgctagc tgctctggcg gcaccgattt cggtagcttc 420
 caaagtgcg ggcgccaccta caacgtgcgt cgttgccagc gcgtgcaaca gccttcgatc 480
 gatggcacc agactttcta ccaatacttc agcgtgagaa atccgaaaaa aggggttggg 540
 aacatttctg gcaccatcac ctttgctaac cacgtaaact actggagaag cagagggatg 600
 aatcttggta accacgatta ccaagtcttc gctactgaag gctacagaag cacgggttct 660
 tctgacctca ccatcagcca aggcgcaagc aacaacggcg gtggcggcag tagctcaagt 720
 gctccatctg ctggggggcg tagcaagaca atcgtcgtgc gggcacgcgg gactaccgga 780
 caagagcaaa tccgtttgcg ggtgaacaac actattgttc agacctggac cttgtccacc 840
 accatgcgcg actacaccgt caacactaac ttggcaggcg ggtcattggt tgaatacttc 900
 aatgacagcg gcaaccgcga cgtccaagtt gattacatca gcgtaaatgg caatgttcgc 960
 caatccgaaa accaaacctt caacaccggg gtctaccaga acggtgcgtg tggcggcggg 1020
 aacggccgga gcgagtggct ccattgcaac ggtgcaatcg ggtacggcga tatctaa 1077

<210> 370
 <211> 358
 <212> PRT
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<221> SIGNAL
 <222> (1)...(27)

<400> 370
 Met Lys Ser Phe Ile Thr Gly Lys Lys Ile Ala Ala Gly Leu Ile Thr
 1 5 10 15
 Ala Ala Ala Leu Ser Ala Ser Met Val Ser Ala Gln Thr Leu Thr Ser
 20 25 30
 Asn Ser Gln Gly Thr His Asp Gly Phe Phe Tyr Ser Phe Trp Lys Asp
 35 40 45
 Ser Gly Asn Ala Ser Met Asn Leu Leu Ala Gly Gly Arg Tyr Gln Ser
 50 55 60
 Ser Trp Asn Thr Gly Thr Asn Asn Trp Val Gly Lys Gly Trp Asn
 65 70 75 80
 Pro Gly Thr Asn Asn Arg Val Ile Asn Tyr Ser Gly Tyr Tyr Gly Val
 85 90 95
 Asp Asn Ser Gln Asn Ser Tyr Val Ala Leu Tyr Gly Trp Thr Arg Asn
 100 105 110
 Pro Leu Val Glu Tyr Tyr Val Ile Glu Ser Tyr Gly Ser Tyr Asn Pro
 115 120 125
 Ala Ser Cys Ser Gly Gly Thr Asp Phe Gly Ser Phe Gln Ser Asp Gly
 130 135 140
 Ala Thr Tyr Asn Val Arg Arg Cys Gln Arg Val Gln Gln Pro Ser Ile
 145 150 155 160
 Asp Gly Thr Gln Thr Phe Tyr Gln Tyr Phe Ser Val Arg Asn Pro Lys
 165 170 175
 Lys Gly Phe Gly Asn Ile Ser Gly Thr Ile Thr Phe Ala Asn His Val

180 185 190
 Asn Tyr Trp Arg Ser Arg Gly Met Asn Leu Gly Asn His Asp Tyr Gln
 195 200 205
 Val Leu Ala Thr Glu Gly Tyr Arg Ser Thr Gly Ser Ser Asp Leu Thr
 210 215 220
 Ile Ser Gln Gly Ala Ser Asn Asn Gly Gly Gly Ser Ser Ser Ser
 225 230 235 240
 Ala Pro Ser Ala Gly Gly Gly Ser Lys Thr Ile Val Val Arg Ala Arg
 245 250 255
 Gly Thr Thr Gly Gln Glu Gln Ile Arg Leu Arg Val Asn Asn Thr Ile
 260 265 270
 Val Gln Thr Trp Thr Leu Ser Thr Thr Met Arg Asp Tyr Thr Val Asn
 275 280 285
 Thr Asn Leu Ala Gly Gly Ser Leu Val Glu Tyr Phe Asn Asp Ser Gly
 290 295 300
 Asn Arg Asp Val Gln Val Asp Tyr Ile Ser Val Asn Gly Asn Val Arg
 305 310 315 320
 Gln Ser Glu Asn Gln Thr Tyr Asn Thr Gly Val Tyr Gln Asn Gly Ala
 325 330 335
 Cys Gly Gly Gly Asn Gly Arg Ser Glu Trp Leu His Cys Asn Gly Ala
 340 345 350
 Ile Gly Tyr Gly Asp Ile
 355

<210> 371
 <211> 1245
 <212> DNA
 <213> Unknown

<220>
 <223> obtained from an environmental sample.

<400> 371
 gtgaccggga tcgcgagaaa aggcgtatgg tccgtgattt ccggaacttt cactgccggg 60
 gattacgatt cctacctgct gtatgtcgaa acacaggacc agggcggcgg acacccgacg 120
 ctgagctttg aaatccggaa cttcagactg acggcaccgg aaggcatcgc tccgccgaag 180
 gcgacagaag aaccggctga cgcggcagag gcgacgcctg ttccggcact gagcgagatt 240
 ccgggcctga aggacgtcta cgcggactac tttgacttcg gcgctgcggc gccgcagtat 300
 gcattcggcc tcggccagac ccagctgcag gacctgatga tcagccagtt cagcatcctg 360
 acccctgaaa acgaactgaa accggacagc gtgcttgatg tccagacgag taaaaaactg 420
 gcggcagaag acgaaaccgc ggtggcgatc aggttgaacg ccgcaacgcc gctgctgaag 480
 ttcgcgcaga agaacggcat caaagtgcac ggccatgtgc tggatatggca cagccagacg 540
 ccggaagctt tcttccatga aggatacgtat accaagaaac cctatgtgac gagagaggtt 600
 atgctcggcc gcctggaaaa ctatatccgt gaagtgtgta cgcagacaga ggaacagttc 660
 ccgggcgtga tcgtcagctg ggacgtcgtg aacgaggcga tcgacgacgg tactcactgg 720
 ctgcggaaga cttccagctg gtacaaagtc gtcggcgagg atttcctgaa cagggctttt 780
 gaatacggca ggaaatacgc cgcggagggc gtgctgtgtg actacaacga ttacagcagc 840
 gcaaattcgg ctaaactgat gggcatcacg aagctgtgta agcagctgat tccagacggg 900
 aatatcgacg gctacggatt ccagatgcac catgacctcg gctggccgag catcgacctt 960
 atggcggcag ctgtgaagca gattgccggc ctggggctga aactgcgcgt cagcgaactg 1020
 gatatcggcg tatccaagaa caatcaggaa aactatgaca aacaggccaa acgctacaag 1080
 gaaatgctga acctgatgct gcagtacgcg gaccagacgg aagccgtgca ggtctggggc 1140
 ctgacggaca acatgagctg gagaaccggc aaataccgcg tgctgttcga cagcgcggca 1200
 aaaccgaaaa aggcgttctt cgcggtgatt gaagccgcag aggaa 1245

<210> 372
 <211> 415
 <212> PRT
 <213> Unknown

<220>
 <223> Obtained from an environmental sample.

<400> 372
 Met Thr Gly Ile Ala Arg Lys Gly Val Trp Ser Val Ile Ser Gly Thr
 1 5 10 15
 Phe Thr Ala Gly Asp Tyr Asp Ser Tyr Leu Leu Tyr Val Glu Thr Gln
 20 25 30
 Asp Gln Gly Gly Gly His Pro Thr Leu Ser Phe Glu Ile Arg Asn Phe

35 40 45
 Arg Leu Thr Ala Pro Glu Gly Ile Ala Pro Pro Lys Ala Thr Glu Glu
 50 55 60
 Pro Ala Asp Ala Ala Glu Ala Thr Pro Val Pro Ala Leu Ser Glu Ile
 65 70 75 80
 Pro Gly Leu Lys Asp Val Tyr Ala Asp Tyr Phe Asp Phe Gly Ala Ala
 85 90 95
 Ala Pro Gln Tyr Ala Phe Gly Leu Gly Gln Thr Gln Leu Gln Asp Leu
 100 105 110
 Met Ile Ser Gln Phe Ser Ile Leu Thr Pro Glu Asn Glu Leu Lys Pro
 115 120 125
 Asp Ser Val Leu Asp Val Gln Thr Ser Lys Lys Leu Ala Ala Glu Asp
 130 135 140
 Glu Thr Ala Val Ala Ile Arg Leu Asn Ala Ala Thr Pro Leu Leu Lys
 145 150 155 160
 Phe Ala Gln Lys Asn Gly Ile Lys Val His Gly His Val Leu Val Trp
 165 170 175
 His Ser Gln Thr Pro Glu Ala Phe Phe His Glu Gly Tyr Asp Thr Lys
 180 185 190
 Lys Pro Tyr Val Thr Arg Glu Val Met Leu Gly Arg Leu Glu Asn Tyr
 195 200 205
 Ile Arg Glu Val Leu Thr Gln Thr Glu Glu Gln Phe Pro Gly Val Ile
 210 215 220
 Val Ser Trp Asp Val Val Asn Glu Ala Ile Asp Asp Gly Thr His Trp
 225 230 235 240
 Leu Arg Lys Thr Ser Ser Trp Tyr Lys Val Val Gly Glu Asp Phe Leu
 245 250 255
 Asn Arg Ala Phe Glu Tyr Ala Arg Lys Tyr Ala Ala Glu Gly Val Leu
 260 265 270
 Leu Tyr Tyr Asn Asp Tyr Ser Thr Ala Asn Ser Ala Lys Leu Met Gly
 275 280 285
 Ile Thr Lys Leu Leu Lys Gln Leu Ile Pro Asp Gly Asn Ile Asp Gly
 290 295 300
 Tyr Gly Phe Gln Met His His Asp Leu Gly Trp Pro Ser Ile Asp Leu
 305 310 315 320
 Met Ala Ala Ala Val Lys Gln Ile Ala Gly Leu Gly Leu Lys Leu Arg
 325 330 335
 Val Ser Glu Leu Asp Ile Gly Val Ser Lys Asn Asn Gln Glu Asn Tyr
 340 345 350
 Asp Lys Gln Ala Lys Arg Tyr Lys Glu Met Leu Asn Leu Met Leu Gln
 355 360 365
 Tyr Ala Asp Gln Thr Glu Ala Val Gln Val Trp Gly Leu Thr Asp Asn
 370 375 380
 Met Ser Trp Arg Thr Gly Lys Tyr Pro Leu Leu Phe Asp Ser Ala Ala
 385 390 395 400
 Lys Pro Lys Lys Ala Phe Phe Ala Val Ile Glu Ala Ala Glu Glu
 405 410 415

<210> 373

<211> 1539

<212> DNA

<213> Unknown

<220>

<223> Obtained from an environmental sample.

<400> 373

ttgattggct	gcgatcatgtc	gccgccggaa	gcgggaagtc	cccgttttga	tcttttaacc	60
cggcacttta	atgtcatcac	cgcggaatac	gccatgaagc	ccgcgtcggt	gcagcgcgaa	120
aagggggtgt	ttacttttga	acaggcggac	atgatggtgg	acgcggtatt	ggagcgggga	180
ctgaagatcc	acggacatac	tctggcctgg	caccagcagt	ctccggagtg	gatgaatcat	240
gaggggattt	cccgggacga	agccgtggaa	aatctcaccg	tccacgccaa	aaccgcggcc	300
gctcatttta	gggggagggt	catatcctgg	gatgtactca	acgaggcgat	cattgacaat	360
ccccccaacc	ccgggggattg	gcgggcatcc	ctcaggcaaa	gcccctggta	caaagccata	420
ggcccggtat	acgtggagct	tgtgttcaag	gcggccaggg	aggcggaccc	ggaggcaaaa	480
ctttattata	acgattacaa	ccttgataac	cggaacaagg	ccctggcggt	ttacaacatg	540
gtcaggggaa	tgaacgaaaa	gaatccgaat	ccgggcggca	ggcccctcat	cgacggcggtg	600
ggcatgcagg	gccattaccg	cctgaatacc	aataccgata	acgtgaggct	gtcgtgggaa	660
cggtttattt	ccctgggggt	cgaggtcagc	atcacggagc	tcgatataca	ggccggttcg	720

gattcaaac	agacagagcg	gcagcgggtg	gaacagggcc	tggtctatgc	cgctttgttt	780
accattttcc	gggaacacgc	ggcaaacata	ggccgggtaa	ctttttgggg	acttgacgac	840
ggggcaagct	ggcgttcgc	ggcgagtccc	tgctctttg	ataaaaacct	caacgcaaaa	900
cctgcctttt	acgcggtcct	ggacccggat	tcctttattg	cggaataacag	cgccctgctg	960
atcaggggaag	cgaagagggg	agaggcttat	tatggtacgc	ctgctttagg	cgccgtccct	1020
gatccccctct	gggacagggc	gccttccctc	ccggtggatc	agtacctcat	ggcctggcag	1080
ggcgcttcgg	gaagggcaaa	agtcctctgg	gacgaaaaaa	atctctatgt	gctggtccgg	1140
gttgaaaacg	cggaaataaa	caaggacagt	tccaacagct	acgaacagga	ttcggtcgaa	1200
atttttattg	atgaggataa	ccggaaaagt	tcctttttca	gggaggatga	cgggcagtac	1260
cgggtcaatt	ttgccaacga	ggcgggcttt	aaccctcgt	ccgccggggc	gggggtttgtt	1320
tcggccgcgc	cgggtgatgg	aaaatcctat	accgttacca	tgaagattcc	ctttaaaaca	1380
atagtccccg	gagcggggac	gcgtatcggg	tttgatgtcc	agatcaacgg	cgcgctcgcc	1440
agggggatac	gggagagcgt	ggcggtatgg	aatgatacca	cgggcaattc	atttcaggat	1500
acctcaggtt	acgggggtact	gcgggttagta	aaaaagtaa			1539

<210> 374

<211> 512

<212> PRT

<213> Unknown

<220>

<223> obtained from an environmental sample.

<400> 374

Met	Ile	Gly	Cys	Val	Met	Ser	Pro	Pro	Glu	Ala	Gly	Ser	Pro	Arg	Phe
1				5					10					15	
Asp	Leu	Leu	Thr	Arg	His	Phe	Asn	Val	Ile	Thr	Ala	Glu	Asn	Ala	Met
			20					25					30		
Lys	Pro	Ala	Ser	Leu	Gln	Arg	Glu	Lys	Gly	Val	Phe	Thr	Phe	Glu	Gln
		35					40					45			
Ala	Asp	Met	Met	Val	Asp	Ala	Val	Leu	Glu	Arg	Gly	Leu	Lys	Ile	His
	50					55					60				
Gly	His	Thr	Leu	Ala	Trp	His	Gln	Gln	Ser	Pro	Glu	Trp	Met	Asn	His
	65				70				75					80	
Glu	Gly	Ile	Ser	Arg	Asp	Glu	Ala	Val	Glu	Asn	Leu	Thr	Val	His	Ala
			85					90						95	
Lys	Thr	Ala	Ala	Ala	His	Phe	Arg	Gly	Arg	Val	Ile	Ser	Trp	Asp	Val
			100					105					110		
Leu	Asn	Glu	Ala	Ile	Ile	Asp	Asn	Pro	Pro	Asn	Pro	Gly	Asp	Trp	Arg
		115					120					125			
Ala	Ser	Leu	Arg	Gln	Ser	Pro	Trp	Tyr	Lys	Ala	Ile	Gly	Pro	Asp	Tyr
	130					135					140				
Val	Glu	Leu	Val	Phe	Lys	Ala	Ala	Arg	Glu	Ala	Asp	Pro	Glu	Ala	Lys
	145				150				155						160
Leu	Tyr	Tyr	Asn	Asp	Tyr	Asn	Leu	Asp	Asn	Arg	Asn	Lys	Ala	Leu	Ala
			165					170						175	
Val	Tyr	Asn	Met	Val	Arg	Glu	Leu	Asn	Glu	Lys	Asn	Pro	Asn	Pro	Gly
		180						185					190		
Gly	Arg	Pro	Leu	Ile	Asp	Gly	Val	Gly	Met	Gln	Gly	His	Tyr	Arg	Leu
	195						200					205			
Asn	Thr	Asn	Thr	Asp	Asn	Val	Arg	Leu	Ser	Leu	Glu	Arg	Phe	Ile	Ser
	210					215					220				
Leu	Gly	Val	Glu	Val	Ser	Ile	Thr	Glu	Leu	Asp	Ile	Gln	Ala	Gly	Ser
	225				230					235					240
Asp	Ser	Asn	Gln	Thr	Glu	Arg	Gln	Arg	Val	Glu	Gln	Gly	Leu	Val	Tyr
			245						250					255	
Ala	Ala	Leu	Phe	Thr	Ile	Phe	Arg	Glu	His	Ala	Ala	Asn	Ile	Gly	Arg
			260					265					270		
Val	Thr	Phe	Trp	Gly	Leu	Asp	Asp	Gly	Ala	Ser	Trp	Arg	Ser	Ala	Ala
		275					280					285			
Ser	Pro	Cys	Leu	Phe	Asp	Lys	Asn	Leu	Asn	Ala	Lys	Pro	Ala	Phe	Tyr
	290					295					300				
Ala	Val	Leu	Asp	Pro	Asp	Ser	Phe	Ile	Ala	Glu	Asn	Ser	Ala	Leu	Leu
	305				310					315					320
Ile	Arg	Glu	Ala	Lys	Glu	Gly	Glu	Ala	Tyr	Tyr	Gly	Thr	Pro	Ala	Leu
			325						330					335	
Gly	Ala	Val	Pro	Asp	Pro	Leu	Trp	Asp	Arg	Ala	Pro	Ser	Leu	Pro	Val
			340					345					350		
Asp	Gln	Tyr	Leu	Met	Ala	Trp	Gln	Gly	Ala	Ser	Gly	Arg	Ala	Lys	Val

```

      355              360              365
Leu Trp Asp Glu Lys Asn Leu Tyr Val Leu Val Arg Val Glu Asn Ala
370      375      380
Glu Ile Asn Lys Asp Ser Ser Asn Ser Tyr Glu Gln Asp Ser Val Glu
385      390      395
Ile Phe Ile Asp Glu Asp Asn Arg Lys Ser Ser Phe Phe Arg Glu Asp
405      410      415
Asp Gly Gln Tyr Arg Val Asn Phe Ala Asn Glu Ala Gly Phe Asn Pro
420      425      430
Ser Ser Ala Gly Ala Gly Phe Val Ser Ala Ala Ala Val Asp Gly Lys
435      440      445
Ser Tyr Thr Val Thr Met Lys Ile Pro Phe Lys Thr Ile Val Pro Gly
450      455      460
Ala Gly Thr Arg Ile Gly Phe Asp Val Gln Ile Asn Gly Ala Ser Ala
465      470      475
Arg Gly Ile Arg Glu Ser Val Ala Val Trp Asn Asp Thr Thr Gly Asn
485      490      495
Ser Phe Gln Asp Thr Ser Gly Tyr Gly Val Leu Arg Leu Val Lys Lys
500      505      510

```

<210> 375
 <211> 570
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetically generated polynucleotide

```

<400> 375
atggccctta tggcttcgac attctactgg cacttggtgga ctgatggtat agggacagta      60
aatgctacca atggatctga tggcaattac agcgtttcat ggtcaaattg cgggaatttt      120
gttggttgga aaggctggac taccggatca gcaactaggg taataaacta taatgcccac      180
gccttttcgg tagtgggtaa tgcttatttg gctctttatg ggtggacgag aaattcactc      240
atagaatatt acgtcgttga tagctggggg acttatagac ctactggaac ttataaaggc      300
actgtgacta gtgatggagg gacttatgac atatacacga ctacacgaac caacgcacct      360
tccattgacg gcaataatac aactttcacc cagttctgga gtgttaggca gtcgaagaga      420
cgattggta ccaacaatac catcaccttt agcaaccatg ttaacgcctg gaagagtaaa      480
ggaatgaatt tggggagtag ttggtcttat cagggtattag caacagaggg ctatcaaagt      540
agtgggtact ctaacgtaac ggtctggtaa                                     570

```

<210> 376
 <211> 189
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetically generated polypeptide

```

<400> 376
Met Ala Leu Met Ala Ser Thr Phe Tyr Trp His Leu Trp Thr Asp Gly
1      5      10      15
Ile Gly Thr Val Asn Ala Thr Asn Gly Ser Asp Gly Asn Tyr Ser Val
20      25      30
Ser Trp Ser Asn Cys Gly Asn Phe Val Val Gly Lys Gly Trp Thr Thr
35      40      45
Gly Ser Ala Thr Arg Val Ile Asn Tyr Asn Ala His Ala Phe Ser Val
50      55      60
Val Gly Asn Ala Tyr Leu Ala Leu Tyr Gly Trp Thr Arg Asn Ser Leu
65      70      75      80
Ile Glu Tyr Tyr Val Asp Ser Trp Gly Thr Tyr Arg Pro Thr Gly
85      90      95
Thr Tyr Lys Gly Thr Val Thr Ser Asp Gly Gly Thr Tyr Asp Ile Tyr
100      105      110
Thr Thr Thr Arg Thr Asn Ala Pro Ser Ile Asp Gly Asn Asn Thr Thr
115      120      125
Phe Thr Gln Phe Trp Ser Val Arg Gln Ser Lys Arg Pro Ile Gly Thr
130      135      140
Asn Asn Thr Ile Thr Phe Ser Asn His Val Asn Ala Trp Lys Ser Lys
145      150      155      160

```

<400> 379
atggccctta tggcttcgac attctactgg cacaattgga ctgatggtat agggacagta 60
Page 291

aatgctacca	atggatctga	tggcaattac	agcgtttcat	gggtcaaattg	cgggaattttt	120
gttggttgta	aaggctggac	taccggatca	gcaactaggg	taataaacta	taatgcccac	180
gccttttcgc	cgggtgggtaa	tgcttatttg	gctctttatg	gggtggacgag	aaattcactc	240
atagaatatt	acgtcgttga	tagctggggg	acttatagac	ctactggaac	ttataaaggc	300
actgtgacta	gtgatggagg	gacttatgac	atatacacga	ctacacgaac	caacgcacct	360
tccattgacg	gcaataatac	aactttcacc	cagttctgga	gtgttaggca	gtcgaagaga	420
ccgattggta	ccaacaatac	catcaccttt	agcaaccatg	ttaacgcctg	gaagagtaaa	480
ggaatgaatt	tggggagtag	ttggctctat	caggtattag	caacagaggg	ctatcaaagt	540
agtgggtact	ctaacgtaac	ggctctggtaa				570

<210> 380

<211> 189

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetically generated polypeptide.

<400> 380

Met	Ala	Leu	Met	Ala	Ser	Thr	Phe	Tyr	Trp	His	Asn	Trp	Thr	Asp	Gly
1				5					10					15	
Ile	Gly	Thr	Val	Asn	Ala	Thr	Asn	Gly	Ser	Asp	Gly	Asn	Tyr	Ser	Val
			20					25					30		
Ser	Trp	Ser	Asn	Cys	Gly	Asn	Phe	Val	Val	Gly	Lys	Gly	Trp	Thr	Thr
			35				40					45			
Gly	Ser	Ala	Thr	Arg	Val	Ile	Asn	Tyr	Asn	Ala	His	Ala	Phe	Ser	Pro
			50			55					60				
Val	Gly	Asn	Ala	Tyr	Leu	Ala	Leu	Tyr	Gly	Trp	Thr	Arg	Asn	Ser	Leu
			65		70				75					80	
Ile	Glu	Tyr	Tyr	Val	Val	Asp	Ser	Trp	Gly	Thr	Tyr	Arg	Pro	Thr	Gly
			85						90					95	
Thr	Tyr	Lys	Gly	Thr	Val	Thr	Ser	Asp	Gly	Gly	Thr	Tyr	Asp	Ile	Tyr
			100					105					110		
Thr	Thr	Thr	Arg	Thr	Asn	Ala	Pro	Ser	Ile	Asp	Gly	Asn	Asn	Thr	Thr
			115				120					125			
Phe	Thr	Gln	Phe	Trp	Ser	Val	Arg	Gln	Ser	Lys	Arg	Pro	Ile	Gly	Thr
			130			135					140				
Asn	Asn	Thr	Ile	Thr	Phe	Ser	Asn	His	Val	Asn	Ala	Trp	Lys	Ser	Lys
			145		150					155				160	
Gly	Met	Asn	Leu	Gly	Ser	Ser	Trp	Ser	Tyr	Gln	Val	Leu	Ala	Thr	Glu
			165					170						175	
Gly	Tyr	Gln	Ser	Ser	Gly	Tyr	Ser	Asn	Val	Thr	Val	Trp			
			180					185							